Vol. XIV APRIL, 1930 No. 2

The Journal of
APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University, Athens, Ohio
WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

H. Meltzer and E. M. Bailor. Sex Differences in Knowledge of Psychology Before and After the First Course 107
A. D. Tendler. A Preliminary Report on a Test for Emotional Insight 122
Nathan Israeli. Variability and Central Tendency in Aesthetic Judgments 137
John A. McGeoch and Marion E. Bunch. Scores in the Pressey X-O Tests of Emotions as Influenced by Courses in Psychology 150
Paul A. Witty. Some Results Secured in a Psycho-Educational Clinic 160
A. Jean Brown. An Enquiry into the Standardization of the Kohs' Block-Design Test 178
Alvhh R. Lager. An Empirical Study of the Effects of Grouping Data in Calculation of R by the Pearson Products Moment Method 182
News and Notes 190
Book Reviews 192
New Books and Pamphlets 195

PUBLISHED BI-MONTHLY BY THE EDITOR
All communications should be addressed to the Editor,
Ohio University, Athens, Ohio
Printed by
WAVERLY PRESS, INC.
Baltimore, Md.

Entered as second class matter June 3, 1924, at the post office at Baltimore, Maryland, under the act of March 3, 1879.
Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922

Copyright 1930, by James P. Porter

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University; W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Department of Education, Baltimore, Md.; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

SEX DIFFERENCES IN KNOWLEDGE OF PSYCHOLOGY BEFORE AND AFTER THE FIRST COURSE

H. MELTZER, PH.D.
Psychologist, Psychiatric Clinic, St. Louis
AND
E. M. BAILOR, PH.D.
Dartmouth College

"Writers, usually male, hold forth on the psychology of woman as if they were dealing with a Platonic universal entity. . . . . They treat phenomena which are peculiarly symptoms of the West at the present time, as if they were the necessary efforts of fixed native impulses of human nature."[1] This quotation from Dewey describes the writings in popular journals. It also characterizes the earlier, more serious studies of sex differences. For example, it applies to "Man and Woman" by Havelock Ellis and the discussion of sex differences in Jastrow's "Character and Temperament." Scientists, evidently are not entirely immune from the stereotypes and preconceptions of the day. In more recent studies, however, the concern is with identifiable forms of realities;[2] with what Miles and Terman describe as "an evaluation of the available experimentally demmonstrated findings which may contribute to an understanding of present developments under existing conditions and may also in part suggest the possibilities of future development under

[1] "Human Nature and Conduct," p. 153.

[2] "Studies in Deceit," by May and Hartshorne, Macmillan, 1928; "Play Interests As Evidence of Sex Differences in Aesthetic Appreciation," by Lehman and Witty, American Journal of Psychology, vol. 40, 1928, p. 449–458; and the article by Miles and Terman (See footnote 3) are examples of this approach.

modified conditions."[3] In this particular instance this means not dialectics for and against the conviction that says in effect "It is a delusion to think you can educate a woman; it just isn't in her" but the report of the objective measurement of the actual accomplishments by men and women of known intellectual capacities as well as socio-economic status when confronted by identical objects of knowledge purposed to serve as stimuli to knowing for a given learning period. The purpose of the present study is to report an investigation so planned which included as part of the investigation, the objective measurement of knowledge of the particular field selected for study, psychology, possessed by the men and women studied before as well as after the learning period. The aspects to be investigated are: (I) sex differences in extent of psychology known before and after the first course; (II) differences in self-improvement; (III) differences in relation to self-improvement and knowledge possessed before taking the course with intelligence, socio-economic status, and age.

The facts reported are of course applicable only for the group studied. The smallness of the group is somewhat compensated for by the more comprehensive knowledge of its make-up and its influences. The intellectual, socio-economic and age composition of the group are reported. The writers do not advocate generalization by extension and ask those who have the effective desire to behave that way to keep in mind the composition of the group, the campus setting, the nature of the course, the means of securing data used—all of which will be presented before the findings are reported.

METHODS, MEASURES, SUBJECTS

A. General procedure

On their first and last appearance in class a group of students registered in the first course in psychology were given the same "final" examination. The examination used was composed of

[3] "Sex differences in the association of ideas," Miles and Terman, American Journal of Psychology, xli, no. 2, 1929, p. 165–206.

one hundred elements set in three different forms: true-false, matching, and multiple choice. The intelligence, socio-economic, age and other information of the individuals studied were obtained sometime during the session by means of tests, questionnaires, and interviews.

B. The setting

A State College on the Pacific Coast located in a fairly conservative town of about six thousand. The enrollment of thirty-seven hundred students is made of roughly twice as many boys as girls. Fraternities and sororities are not lacking in numbers. Most of the students come from rural communities or small towns. The women boast of one of the best equipped gymnasiums in the country and of superior possibilities for training in physical education. They have many other opportunities to make "points" for participation in extra-curricular activities. But in spite of all this, the dominant customs and traditions are largely man-made. Men do the sensational "fighting for the honor of the college." And in general women have their lives more regulated by the authorities, have many less socially approved opportunities for emotional outlets in terms of effective space-time possibilities and as a possible sequence more time for study, for developing habits of study which in time by the process described by Dewey in "Human Nature and Conduct" develop into greater and more effective desires to do well in school subjects.

C. Measuring instruments

A more detailed description of measuring instruments used has already been published[4] and will here be hardly more than listed: Intelligence was measured by the Otis Self Administering Test of Mental Ability, Higher, Form A; Socio-Economic Status was determined by the use of the Chapman-Sims

[4] Measuring instruments used are described more fully in article by the same authors, "Extent and Nature of Psychology Students Know Before and After the First Course," in Journal of Genetic Psychology, forthcoming issue.

Scale; Psychology grade was determined entirely by six objective tests covering all aspects of subject matter of the course and set in a variety of forms.

D. The nature of the course

The nature of the course is described in full in the introduction of the book[5] which is the outcome of consciously planned efforts to psychologize the instruction of the first course in psychology. The lessons were purposed to serve as a guide to the intelligent use of many books as source materials for active learning. Objective tests with norms were included in these lessons to serve in the motivation of learning. Another series of objective tests was used for measurement. The discussion group method of teaching was used after the first week of the course. The size of discussion groups was 6 to 10. Leaders were changed weekly.

E. The subjects—intellectual and socio-economic status of men and women subjects

Fifty-six students took the final examination on the first giving. Of these 28 were men and 28, women. Of the 67 who took the same examination on their last appearance for the course 31 were women and 36, men.

The differences in distribution of intelligence, chronological age and socio-economic status in the men and women studied can be summarized as follows: In intelligence the differences are insignificant. Analysis of their Otis scores indicate this fact:

SEX	MEDIAN	Q	NUMBER OF INDIVIDUALS
Men	55.75	7.25	34
Women	55.50	8.00	32

[5] "Developed Lessons in Psychology, Including Objective Tests with Norms," by H. Meltzer and E. M. Bailor, Harcourt, Brace and Company, 1929.

Differences in age distribution of the two sexes are small:

SEX	MEDIAN	Q	NUMBER OF INDIVIDUALS
Men	22	2	31
Women	21	1	30

The largest difference is in the social background. As a group the girls rate themselves on the Chapman-Sims Score Card two levels higher than the boys—the boys' median is on the fifth level; the girls' on the seventh level of socio-economic status.

SEX	MEDIAN	Q	NUMBER OF INDIVIDUALS
Men	50	2	33
Women	53	3.5	28

At least part of this difference may however be explained by what seemed to be a more frequent tendency of more of the girls to overestimate the number of things in their home that are indicative of richness. That this difference in level of socio-economic status as far as influencing self-improvement or success in course is concerned is not significant is indicated by the extent of the correlation between socio-economic status and gain in knowledge of psychology reported in a previous article.[6] ($r = -.017 \pm .078$).

I. SEX DIFFERENCES IN EXTENT OF PSYCHOLOGY KNOWN BEFORE AND AFTER THE FIRST COURSE

The extent of knowledge of psychology students possessed before and after the course was determined by an analysis of the scores made by them on the one-hundred element final examination taken in the two respective times, i.e., the first

[6] Meltzer and Bailor, "Extent and Nature of Psychology Students Know Before and After the First Course," Journal of Genetic Psychology, forthcoming issue.

and last appearances in class. The tests were scored according to general practice—right minus wrong for true-false and number of right for matching and multiple choice. The total score was, of course, the sum total of the scores on the three different parts of the examination. The highest possible score was thus 100. The most concise and direct way of indicating the sex differences in knowledge found is probably in terms of measures of central tendency and variability before and after. Such measures and others are reported in table 1.

A more detailed picture of the distributions of scores before and after can be obtained if the two distributions are given

TABLE 1

Showing differences in mean scores before and after

SEX	TIME	NUMBER OF INDIVIDUALS	MEAN	S.D.	HIGHEST SCORE	LOWEST SCORE
Men	Before	28	24.2	9.68	49	12
Women	Before	28	24.7	12.14	61	1
Men	After	36	62.6	13.09	82	33
Women	After	31	72.3	10.31	89	35

in full and plotted on the same base. Figure 1 shows such a picture for men and figure 2 does the same for women.

An analysis of the foregoing table and graphs reveals the following facts: The amount of knowledge of psychology that the women in this group knew before taking the course was insignificantly higher than that known by the men in the group. But that, after the course, the mean score for the women is significantly higher than it is for the men. A comparison of S.D.'s as well as a comparison of the ratio of poorest to best in the respective groups indicate less variability in the women than in the men on the test taken before the course but more variability in the women on the test taken after the course. The range for men was 25 points less after the course than before; the range for women was 14 points more on the test taken after the course. No score of either the men or

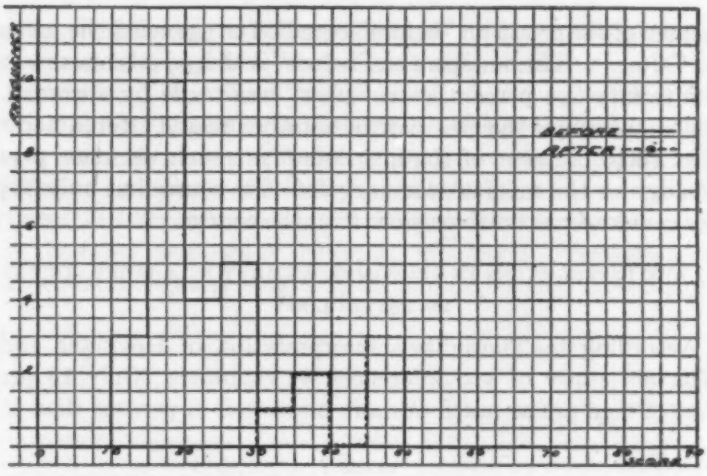

FIG. 1. SHOWING THE DISTRIBUTION OF SCORES MADE BY COLLEGE STUDENTS ON SAME PSYCHOLOGY TEST TAKEN BEFORE AND AFTER THE COURSE: MEN

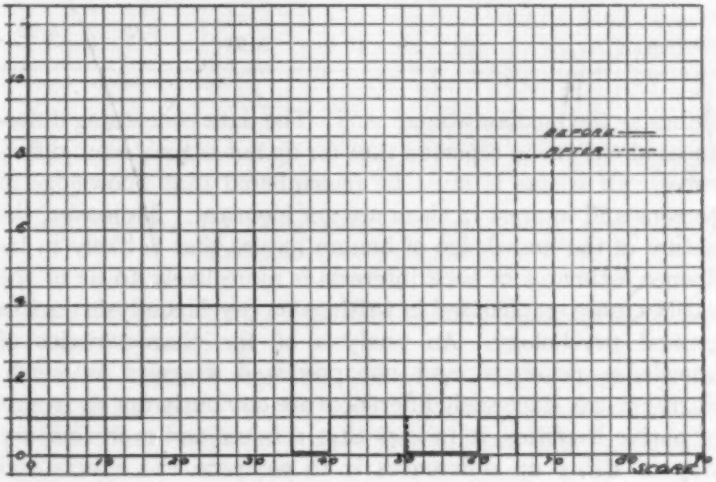

FIG. 2. SHOWING THE DISTRIBUTION OF SCORES MADE BY COLLEGE STUDENTS ON THE SAME PSYCHOLOGY TEST TAKEN BEFORE AND AFTER THE COURSE: WOMEN

women of the group on "before" equalled or exceeded the mean on "after." The one score on "before" which equals or exceeds the twenty-five percentile on "after" was made by a woman. A comparison of the two distributions before and after indicate that 46.15 per cent of the men knew as much or more psychology as the median woman before the course but only 33⅓ per cent knew as much or more after the course.

II. SEX DIFFERENCES IN SELF-IMPROVEMENT IN A FIRST COURSE IN PSYCHOLOGY

According to Thorndike's interpretation of Heymans and Wiersma study of sex differences by estimates "Only 15 per cent of men are as much or more interested in persons than in

TABLE 2

Sex differences in mean gains in self improvement

SEX	MEAN GAIN	S.D.	DIFFERENCE (D) IN FAVOR OF WOMEN	S.D.
Men	38.4	3.42	9.2	5.02
Women	47.6	3.68		

things as the median woman is."[7] The data concerning differences in our group reported thus far do not suggest such differences. Of course it must be remembered that knowledge of persons and interest in them are not identical. Moreover, the nature of the knowledge of human nature that is likely to influence estimates of such interests is more likely to be of the variety more emphasized in writings on psychoanalysis than in general psychology. An analysis of the characteristics of knowledge possessed and not possessed suggests as much. Furthermore, since such differences are obviously not the result of native impulses, it follows that a fairer and more unitary way of sizing up differences is to compare the efforts and achievements of the two sexes when confronted by similar

[7] E. L. Thorndike, Educational Psychology Briefer Course, p. 348.

stimuli to knowing. The results found on test taken after the course already reported suggest such differences.

A study of individual gains made from before to after will reveal such differences in a more prehensive manner. The gain in points on the 100 element objective examination was taken as a measure of self-improvement. Unlike success in a course, self-improvement is not dependent on the scores made by the other members of the class. The point differences were not equalized by any weighting techniques because according to

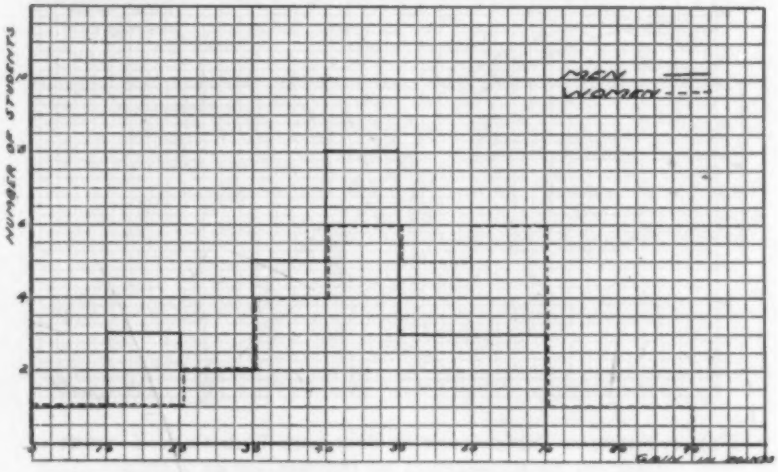

FIG. 3. SHOWING SEX DIFFERENCES IN DISTRIBUTION OF GAIN IN KNOWLEDGE OF PSYCHOLOGY

evidence reported by some authorities no increase in accuracy is obtained by so doing.[8]

How the gains of men compare with those of women is indicated in figure 3.

A statistical comparison of the data in figure 3 indicates that

[8] C. F. Ruch and Stoddard, "Tests and Measurements in High School Instruction," p. 336.

only 24 per cent of the men equal or exceed the gain in knowledge by the median women.

The extent of differences in mean gains is indicated in table 2. The difference D is thus 1.83 of the S.D. difference. This means the chances are 96 out of 100 that the difference is real.[9]

The extent of sex differences in terms of percentage of students who make various amounts of self-improvement is

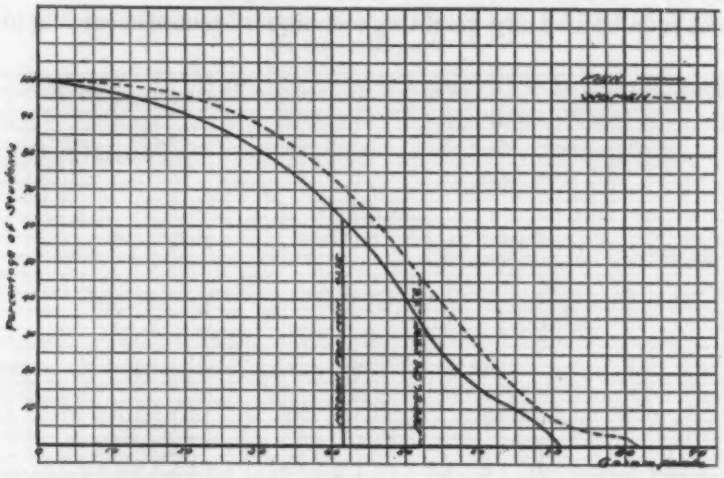

FIG. 4. SHOWING PERCENTAGE OF STUDENTS WHO GAIN VARIOUS AMOUNTS OF KNOWLEDGE OF PSYCHOLOGY IN A ONE TERM LEARNING PERIOD

indicated in figure 4. The median gain made by the women is 52 while that for the men is 41.75.

III. DIFFERENCES IN RELATION OF SELF-IMPROVEMENT AND KNOWLEDGE POSSESSED BEFORE TAKING THE COURSE WITH INTELLIGENCE, SOCIO-ECONOMIC STATUS AND AGE

The extent of sex differences in the relation of traits considered will be first indicated by the correlation method. This

[9] Cf. Table XIV, Garrett's "Statistics in Psychology and Education."

will be supplemented by an arithmetical analysis of gains of men and women from extreme ends of distribution on trait or factor considered. The Spearman Rank Difference Method was used in determining P's and these were converted to r's by the use of tables in Garrett. In table 3 are reported the correlations of self improvement with other factors and in table 4 the

TABLE 3

Showing differences in extent of correlation between self-improvement and other factors

FACTORS	BOYS				GIRLS			
	N	ρ	r	PEr	N	ρ	r	PEr
Self Improvement and								
1. Intelligence (Otis score).	22	−.05	−.05	±.14	25	.10	.10	±.14
2. Socio-Economic Status (Chapman-Sims score).	20	−.06	−.06	±.14	23	.06	.06	±.14
3. Age	20	.22	.23	±.14	24	−.07	−.07	±.14

TABLE 4

Showing differences in extent of correlation between knowledge possessed before course and other factors

FACTORS	BOYS				GIRLS			
	N	ρ	r	PEr	N	ρ	r	PEr
Score "Before" and								
1. Intelligence	23	.48	.49	±.11	25	.07	.07	±.13
2. Socio-Economic Status	21	.12	.13	±.13	23	.27	.28	±.13
3. Age	21	.12	.13	±.13	24	−.06	−.06	±.14

correlation of these factors with score on test taken before taking the course.

The PE of the r's in table 3 are large and hence the only statement which can be made with some safety at this point in the study is that the correlations of these factors with self-improvement are not significant.

The largest difference in these correlations is that between knowledge possessed before and intelligence. What explains

the larger correlation for men is not apparent. One explanation may be the larger amount of venturesomeness of men on the before test. The higher correlation for men on this basis may even be indicative of nothing more than that the more intelligent men are better guessers. For the women the correlation following this inference is lower because of the smaller extent of venturesomeness in the women of the group. The higher correlation of score on before test with socio-economic status for women suggests that women profit more from socio-economic status than do men.

TABLE 5

Showing differences in average gains of men and women from extreme levels of intelligence, socio-economic status and age

FACTOR	AVERAGE GAIN		DIFFERENCES IN FAVOR OF WOMEN
	Men	Women	
Superior Intelligence (above 50 Otis).....	33.62	53.13	19.51
Inferior Intelligence (below 48 Otis).....	34.20	46.57	12.37
Older Students (CA-25 and above).....	37.33	72.50	39.17
Younger Students (CA-20 and below).....	36.25	53.42	17.17
High Level Socio-Economic Status (above 53 Chapman-Sims Score).....	39.40	44.00	4.60
Low Level Socio-Economic Status (below 48 Chapman-Sims Score).....	30.30	50.00	19.67

Further light on the nature and extent of the relation of self-improvement to factors studied may be obtained by analyzing the extent of self-improvement in terms of average number points gained of men and women from extreme levels of intelligence, socio-economic status and age. The results of such an analysis are shown in table 5.

The data of the foregoing table indicate that the women in the group make larger gains than men regardless of extreme level of intelligence, socio-economic status or age. It also indicates, however, that the difference in gains in favor of women is larger by 7.14 per cent in the individuals of superior intelligence than in those of inferior. This suggests lack of sufficient

attention and effort of male students and the greater need for the stimulation of brighter men students than women. Results also indicate a 20 per cent larger difference in favor of women in the group of older students than in the younger students. In socio-economic status the picture is reversed and we find 15.07 per cent less difference in individuals coming from higher levels than those from the lower. In interpreting these results the small number of cases and the possible existence of spurious factors must of course be kept in mind. It is not the intentions of the writers to argue that these differences will be found in other groups or in other circumstances; they are applicable only to the group studied. The differences have been reported in this manner as much to illustrate the use of supplementary common sense methods to throw further light on the nature of statistical results—in this case the correlation of self-improvement with intelligence, socio-economic status and age—as to indicate quantitative differences.

SUMMARY

This study reports the differences in the self-improvement of a group of men and women of known intellectual capacities and socio-economic status in the first course in psychology, the nature of which is described. Knowledge before and after the course was measured by the same 100 element objective examination given on the first and last appearance of the students in class. Intelligence, socio-economic status, age and success in a course were secured by means of tests, scales, interviews and questionnaires given during the session.

On intelligence and age the distribution of the two sexes show no significant differences; on socio-economic status the median for women are two levels higher than the median for men on the Chapman-Sims scale. The results found may be summarized as follows:

I. Analysis of Differences in the Extent of Knowledge possessed by the men and women in the group indicate that the differences are very small indeed before the course but sub-

stantially in favor of the women of the group after course. The more general relevant facts are:

1. The mean for women before the course was 24.7; for men it was 24.2. The mean for women after the course was 72.3; for men it was 62.6.

2. On the test taken before the course 46.15 per cent of the men equal or exceed the score of the median women. On the test taken after, only 33⅓ per cent of the men equal or exceed the median score of the women in the group.

3. A comparison of S.D.'s as well as ratios of the poorest to the best indicate less variability in women before the course but more after the course.

II. Analysis of differences in self-improvement as measured by gain in knowledge of psychology from "before" test to "after" indicates significantly larger gains for women. The more general relevant facts are:

1. The mean difference in gains in favor of women is 9.2. That this is statistically significant is indicated by the fact that $\frac{D}{S.D.} = 1.83$. This means the chances are 96 out of 100 that the difference is real.

2. On gains made only 24 per cent of the men equal or exceed the median gain for women.

III. A study of sex differences in self-improvement in relation to socio-economic status, age, and intelligence made by two approaches—correlation and comparison of gains on extreme levels—brings out some interesting results.

1. Correlation studies indicate low and insignificant r's with intelligence, socio-economic status and age for the two sexes. There are fair-sized differences in the correlations between score on before test and intelligence ($.49 \pm .11$ for men; $.07 \pm .13$ for women). The explanation offered for this difference in favor of men is in terms of implications of the larger extent of venturesomeness found in men. The higher correlation of score on "before" test with socio-economic status ($.13 \pm .13$ for men; $.28 \pm .13$ for women) suggests that women profit more from socio-economic status than do men.

2. Analysis of gains made by men and women from extreme levels of factors considered show larger gains for women from all levels. But they also show that the difference in gains in favor of women is greater by 7.14 per cent in individuals from superior levels of intelligence than in those of inferior; is 20 per cent larger in favor of women in group of older students than in younger; but is 15.07 per cent less in individuals coming from higher socio-economic status than those coming from lower. These facts suggest that brighter men are working up to their capacities less so than women and also less than dull men; that older men of the group can stand more motivation than older women; and that high socio-economic status does not make for higher correlation of ability and effort in either sex.

A PRELIMINARY REPORT ON A TEST FOR EMOTIONAL INSIGHT

A. D. TENDLER

College of the City of New York

INTRODUCTION

Techniques that have been fruitful in the study of intellectual functions have been applied to developing tests for the emotions without adequate consideration as to how suitable the methods may be for emotional functions. Devices that are provocative of discrimination are excellent for the diagnosis of thought reactions but may be wholly unsuited for the diagnosis of emotional responsiveness. When the attempt is to evoke trends, fixed attitudes, attachments to persons, conflicting desires, satisfactions and annoyances, such devices may serve only to dispel rather than evoke the desired responses.

CRITERIA FOR EMOTIONAL TESTING

It is possible to hazard criteria for test methods in the field of emotional behavior. These may well be:

1. The use of such devices that directly evoke emotional response.
2. That allow for free response.
3. That do not provoke discrimination or choice.

From such criteria it appears that the approach of the physiological-psychologist or the behaviorist variant ranks fairly high in adequacy of method. However, the importance of the verbal emotional association, the richness and variety of emotional states and attitudes of which language gives evidence, seem to demand a verbal approach as well. It is possible to work out a test that will be in keeping with the criteria suggested.

A brief analysis of the verbal tests now in use for emotional diagnosis will help to pave the way for the new test presented.

1. The free association test technique satisfies the free response criterion but falls short in adequately arousing a particular emotional set. A single word as a stimulus has such a multiplicity of associations that it may often hit wide of the mark. In this way the directness factor is violated. Discrimination is probably minimal.

Name.....Age.....Grade.....Time.....Score.....

Sample

I eat when—I feel hungry
I sleep when—I feel tired
I read—books

1. My hero is.....
2. I get angry when.....
3. I feel happy when.....
4. I love.....
5. I hate.....
6. I feel hurt when.....
7. I worry over.....
8. I make believe.....
9. I feel sorry when.....
10. I brag about.....
11. I feel proud when.....
12. I have a grudge against.....
13. I become stubborn when.....
14. I pity.....
15. I feel ashamed when.....
16. I am afraid when.....
17. I like to.....
18. I become disgusted with.....
19. I tell lies when.....
20. I wish for.....

FIG. 1. TEST FOR EMOTIONAL INSIGHT

Devised by A. D. Tendler

---

2. The Yes-No method as in the Woodworth Personal Data blank involves a degree of discrimination in judging as to whether a particular trait is characteristic of oneself. It does perhaps directly evoke a motivation attitude as to whether one is to admit the possession of more or less easily recognized unfavorable traits. The response is fixed, not allowing for individual variation in degree of the trait.

3. Multiple choice tests, cross out, order of merit, scale evaluations, all involve the discrimination factor to too great an extent and hence are more likely to provoke thought than emotional reactions.

DESCRIPTION OF EMOTIONAL INSIGHT TEST

The new test presented by the writer is designed as a test for emotional insight, insight being thought of as knowledge of the situations under which particular emotional responses are made. The technique made use of is the completion test. The stimuli in incomplete sentences are given in such form that they directly arouse a particular emotional set and yet allow for all that is implied in free response. By the arousal of particular emotional sets discriminative responses are eliminated. Figure 1 is a copy of the test.

In giving the test, the sample is explained. The subject is told to complete the sentences on the blank, to respond promptly, to expose a single stimulus at a time, and to answer all the items on the blank. A sheet of paper is given to serve as a screen for the stimuli.

The stimuli include a major share of important emotional states: admiration, anger, happiness, love, hate, self abasement worried state, compensatory make believe, regret, boastfulness, pride, grudge formation, negativisim, pity, shame, fear, interest, disgust, evasion desire. The selection of emotional states is on an empirical basis. There may be other states it may be desirable to include. There may be similar elements in various states, the total set having however, a distinguishing reference. Undoubtedly the stimuli require further analysis. The statement of the stimuli is, however, in terms commonly used and commonly understood so that they have significance for the individual responding. Again the particular mode of statement was chosen with an eye towards its use with a wide range of subjects—children, adults of varying degrees of education, age, experience and life adjustment.

ANALYSIS OF RESPONSES

A. General

To indicate the way subjects behave on the test table 1 is presented. Table 1 was obtained as follows: The Emotional

TABLE 1

Responses of subjects

STIMULI	SUBJECT A	SUBJECT B	SUBJECT C	SUBJECT D	SUBJECT E	SUBJECT F
1. My hero is	the ideal man	a fine man	Cyrano de Bergerac	a tall man, dark handsome intelligent man	Abr. Lincoln	Abr. Lincoln
2. I get angry when	thwarted	I am annoyed	I am abused	I quarrel	I am forced to waste time needlesaly	I am mistreate
3. I feel happy when	in the company of others	I am among friends	things go my way	I am having a good time	fortune smiles	I am interest in things
4. I love	people	my family	a person with sharp wits		most members of my immediate family and a few others	to do social se vice work
5. I hate	vulgarity and loudness	my foes	suggestive jokes	(when I am hurt by someone) H. H.	selfishness	fussy people
6. I feel hurt when	insulted	I am reprimanded	I am disbelieved	someone is unkind	I am slighted by those I love	I am not appr ciated
7. I worry over	school	nothing	hardly anything except family budget	my father	my mother's health	little things
8. I make believe	about pleasant situations	that I am always satisfied		that I'm popular and good looking	I am contented all the time	that I do n worry
9. I feel sorry when	others are hurt	I hurt someone	I tease someone	someone is dead or sick	people hurt each other by nasty words	I hurt people
10. I brag about	my sewing	nothing	my ability to take care of the radio	my friends	my ability to keep answers to questions to myself	my friends
11. I feel proud when	I succeed	I am praised	I get a good mark on my paper	about my position	I succeed in being unselfish	people praise
12. I have a grudge against	no one	no one	no one	people who are affected	people who are always late	a girl who u to work with

TABLE 1

Responses of subjects

SUBJECT C	SUBJECT D	SUBJECT E	SUBJECT F	S
Cyrano de Bergerac	a tall man, dark handsome intelligent man	Abr. Lincoln	Abr. Lincoln	Osca
I am abused	I quarrel	I am forced to waste time needlessly	I am mistreated	some thw
things go my way	I am having a good time	fortune smiles	I am interested in things	I am
a person with sharp wits		most members of my immediate family and a few others	to do social service work	Osca
suggestive jokes	(when I am hurt by someone) H. H.	selfishness	fussy people	nobo tic
I am disbelieved	someone is unkind	I am slighted by those I love	I am not appreciated	I am
hardly anything except family budget	my father	my mother's health	little things	littl
	that I'm popular and good looking	I am contented all the time	that I do not worry	I fin
I tease someone	someone is dead or sick	people hurt each other by nasty words	I hurt people	I see
my ability to take care of the radio	my friends	my ability to keep answers to questions to myself	my friends	my (wh goo
I get a good mark on my paper	about my position	I succeed in being unselfish	people praise me	I do wo
no one	people who are affected	people who are always late	a girl who used to work with me	thos acid wr me

JECT F	SUBJECT G	SUBJECT H	SUBJECT I	SUBJECT J
ncoln	Oscar	Lindbergh	handsome	one who never takes advantage over anyone
nistreated	someone thwarts me	I miss a train	I am thwarted	I see people wrangle constantly
interested ngs	I am pleased	I am reading	something happens that is nice	people I like are with me
ocial serork	Oscar	dogs	my father	people who are agreeably sarcastic
eople	nobody in particular	monkeys	very few people	to see a person taken advantage of
ot appre-	I am insulted	I make a faux pas	people laugh at me	I am accused of some petty thought which I never entertain
ings	little things	my dog	my feelings of inferiority	anything that is not connected with school
do not	I finished school	I am older than I am	I am wealthy	I'm happy when I feel most like bawling
people	I see tragedy	I hurt someone	I have done something that hurts someone else	moments of happiness are over
nds	my marks (when they are good)	my dog	nothing	my free and easy manner with men but my friends know me
praise me	I do something worthwhile	I get decent marks	I've accomplished something in school	someone I like achieves something
who used k with me	those who consciously do wrong against me	an acquaintance	nobody	the male sex in general and against one in particular

STIMULI	SUBJECT A	SUBJECT B	
13. I become stubborn when	I don't get what want	I am contradicted	I a o
14. I pity	those who do not succeed	those who are unfortunate	th
15. I feel ashamed when	failure happens	I have done something wrong	I m
16. I am afraid when	in dangerous positions	I am alone	th b sp
17. I like to	draw	read a grand deal	aw
18. I became disgusted with	insolent people	uninteresting things	ha
19. I tell lies when	in situations unfavorable to me	I have to	th a
20. I wish for	happiness and success	happiness	ha

TABLE 1—*Concluded*

BJECT B	SUBJECT C	SUBJECT D	SUBJECT E	SUBJECT F	SUBJE
contra- d	I am opposed by one not certain	I am angry or disappointed	ever anyone says anything over positive	I believe I am right	I am th
who are un- nate	the poor and sick	invalids	the world for senseless wars it engages in	poor and sick	unfortun
ave done thing g	I have offended mother	I'm caught doing something wrong	I am caught in an untruth	I am rude to a mentally deficient person	I do a wrong
lone	the electric fuse burns out in sparks	I see a mouse	people ask me personal questions	I am alone in the dark	somethin wrong
gread deal	swim	see shows	talk to friends	play tennis	read
eresting s	habitual liars	improper jokes	people who talk food	unpleasant people	myself s times
to	they won't harm anybody	I'm with H. H.	it seems to spare the people I tell them to	it will hurt no one and help someone	it is abs necessar (white o
ness	happiness	happiness and love	leisure	health and money	happines

.—*Concluded*

UBJECT E	SUBJECT F	SUBJECT G	SUBJECT H	SUBJECT I	SUBJECT J
anyone says thing over tive	I believe I am right	I am thwarted	I am commanded sharply	the other person becomes insistent	I feel that I have been too easy going
world for eless wars it ges in	poor and sick	unfortunates	very young children	the blind	people with inferiority complexes
caught in an uth	I am rude to a mentally deficient person	I do anything wrong	I've lied	I've done something agninst the advice of other people and it turns out wrong	one discovers that underneath my assumed gaiety there is an air of misery—sometimes
e ask me onal ques- s	I am alone in the dark	something goes wrong	I'm alone in the house	I'm on a high place	I see lightning
o friends	play tennis	read	read	dance	go to the theatre
e who talk	unpleasant people	myself sometimes	gossip	persistent stupidity when I have tried to teach the right thing	conceited fool males
ns to spare eople I tell to	it will hurt no one and help someone	it is absolutely necessary (white ones)	it's necessary	I'm afraid to be discovered	I feel that my imagination is stagnating
	health and money	happiness	a new house	many things	unalloyed happiness for at least a year and then it doesn't matter what happens to me

Insight Test blanks of 250 college girls, 75 per cent of whom fall at the ages 18, 19, 20, who had also been given a form of the Woodworth Personal Data blank. These were arranged in an order of increasing Personal Data unfavorable scores—from 0 to 41 out of 70 items. Every 25th Emotional Insight Test blank was drawn out and constituted the series of 10 presented in table 1 for demonstrating general features of response and method of individual analysis. The first column gives the stimuli. The succeeding columns give the responses of individuals A, B, C, D, etc. The individuals are presented in an increasing order of unfavorable scores on the Personal Data blank.

In analyzing table 1, observations may be made in horizontal or vertical directions. The horizontal examination, reading across, should reveal features that are likely to appear when any number of people are given the test. Vertically, reading down the columns, items of significance for individual analysis are indicated.

The following features may be observed in an horizontal examination of table 1.

a. The same stimulus evokes different responses from different individuals.

For example:

Stimulus 1, my hero is, evokes: 4 conventional characterizations (the ideal man; a fine man; tall, dark, handsome, intelligent man; handsome), 3 literary and historical characters (Cyrano, Lincoln), 1 popular character (Lindbergh), 1 personal acquaintance (Oscar), 1 emotional characterization (one who never takes advantage over anybody).

Stimulus 2, I get angry when, evokes: 6 defensive reactions (thwarted, mistreated, annoyed, abused), 1 aggressive (I quarrel), 2 objective interference (forced to waste time, miss a train), 1 emotional setting (see people wrangle constantly).

The same feature is seen in the responses to all of the stimuli. The Kent-Rosanoff Free Association Test results in commu-

nity of response seem to hold in this test too and suggests a possible mode of standardization.

b. Individuals differ in the associative flow of response. Subject A uses 54 words in responding to all the stimuli, B uses 58 words, C 78, D 74, E 108, F 82, G 62, H 57, I 89, J 146—a range of 54 to 146 words. This feature may be of significance as expressing a feeling of need for completeness of response. It may be allied to a tendency observed in the free association test for subjects not to be satisfied with single word responses and to add additional comments. Jung has treated it as an indication of the hysteric personality. Janet's characterization of the hysteric's tendency to carry behavior to extreme degrees of completion may fit in with this feature.

c. Responses are indicative of satisfactions and annoyances. Practically all of the responses are subject to this category. The satisfactions and annoyances produced may be seen to be dependent upon the stimuli. Stimuli 1, 3, 4, 8, 10, 11, 17, 20—my hero is, I feel happy when, I love, I make believe, I brag about, I feel proud when, I like to, I wish for, are provocative of satisfactions. Stimuli 2, 5, 6, 7, 9, 12, 13, 14, 15, 16, 18, 19—I get angry when, I hate, I feel hurt when, I worry over, I feel sorry when, I have a grudge against, I become stubborn when, I pity, I feel ashamed when, I am afraid when, I become disgusted with, I tell lies when, these provoke annoyances. The responses given are at times specific, at times general. What significance to attach to specific or general responses will require specific study. Factors that suggest themselves in this connection are: education, maturity, insight, evasiveness.

d. Responses are indicative of fears, aversions, likes, interests, and attachments. The responses produced have reference to persons, activities, social situations, intellectual matters, animals, etc. Instances of these may be seen in liking to draw, to talk to friends, to read; becoming disgusted with improper jokes, people talking about food; being afraid when alone, when I see a mouse; pitying the world for the senseless wars it engages in; love to do social service work, love for family, for father, for specific individual, for person with sharp wits.

e. Responses may have an ego or social reference. Each in turn may have a positive or negative phase. Thus an ego positive response has an assertive reference, as, I feel proud when I succeed, I feel happy when things go my way. An ego negative response has a self depreciative reference, as, I feel proud when I succeed in being unselfish, I become disgusted with

myself, I worry over my feelings of inferiority, I feel hurt when I am not appreciated.

A social positive response is one that suggests interest and feeling for others, as, I love people, I worry over my mother's health, I pity the poor and sick, I like to talk to friends, I feel sorry when others are hurt, I tell lies when it seems to spare the people I tell them to.

A social negative response is one that is suggestive of dislike, fault finding, not getting on with people; as, I get angry when I quarrel, I hate fussy people, I feel hurt when someone is unkind, I become disgusted with insolent people, I have a grudge against people who are affected, I have a grudge against a girl who used to work with me, I hate my foes.

The general character of the responses have been indicated. The features educed are by no means exhaustive. The important point is that the responses have a verisimilitude about them. They do give the psychologist insight into matters that are of emotional concern to these young people, regardless of any classificatory scheme.

B. Individual

Now as to the possibilities for individual analysis. The responses on the Emotional Insight Test do give a picture of the individual and the features previously outlined are readily found. In addition, examination of individual test papers brings to light a significant feature—the tendency to respond in similar fashion to several different stimuli—a perseveration of response. It suggests itself as significant of an emotional trend. Samples of interpretation of individual test papers are given. If the reader will follow by reading down the columns of table 1 under the appropriate letter, he will have at his disposal the data from which the analysis is drawn.

Case A. Considerable perseveration on success and failure is shown. Six of the responses deal with this content. Stimuli 11 and 20—I feel proud when I succeed, I wish for success and happiness—deal with the success factor. Stimuli 2, 13, 14, 15—I get angry when thwarted, I become stubborn when I don't get what I want, I pity those who do not succeed, I feel ashamed when failure happens—deal with the failure factor. The social

motive, positive phase, also shows perseveration. Responses 3, 4, 9 and 12 deal with this tendency—I feel happy when in the company of others, I love people, I feel sorry when others are hurt, I have a grudge against no one. Four responses deal with this feature. An ego negative state of affairs is seen in response to stimuli 6 and 18—I feel hurt when insulted, I become disgusted with insolent people. Response to stimulus 5—I hate vulgarity and loudness—may tie up with this ego negative state of affairs. Worry is indicated in response to stimulus 7—I worry over school; fear is seen in response 16—I am afraid when in dangerous positions, its general character suggests an evasion factor. Possible balancing factors or tension relieving activities are seen in responses 10, 17, 8 and 20—I brag about my sewing, I like to draw, I make believe about pleasant situations, I wish for happiness. An indication of evasive behavior is seen in response 19—I tell lies when in situations unfavorable to me.

Comment: A picture of a success obsessed and thwarted individual, socially inclined, inadequate ego sentiment. Specific worry over school; possible balancing activities are sewing and drawing.

Case D presents a different sort of picture. A social positive response is 10—I brag about my friends. Social negative indications are seen in responses 2, 6, 8, 12—I get angry when I quarrel, I feel hurt when someone is unkind, I make believe that I am popular, I have a grudge against people who are affected. Indications of an ego negative attitude are seen in responses 8, 13, 15—I make believe that I am good looking, I become stubborn when angry or disappointed, I feel ashamed when caught doing something wrong. A perseveration appears with respect to sickness in responses 9 and 14—I feel sorry when someone is dead or sick, I pity invalids. Response 7—I worry over my father, may be connected with this. Another perseveration is with respect to an individual H. H. Responses 5 and 19 deal with this—I hate H. H., I tell lies when I'm with H. H. Responses 1 and 20 may be connected with this situation—My hero is a tall man, dark, handsome, intelligent man, and I wish for happiness and love. A fear is seen in response 16—I am

afraid when I see a mouse. A sex attitude is seen in response 18—I become disgusted with improper jokes. Possible balancing or tension relieving activities are seen in responses 17 and 11—I like to see shows, I feel proud about my position.

Comment: Social negative and ego negative attitudes, special concern over father and sickness, tension provoking relation with an individual H. H., sensitivity to sex references, aversion to mice, tension relieving activities: shows and work.

Case G. Considerable stress is shown on ego negative and social negative responses. Responses 6, 15, 18, 19 and 2, 13, 12 bring these out—I feel hurt when insulted, I feel ashamed when I do anything wrong, I become disgusted with myself sometimes, I tell lies when it is absolutely necessary, white ones (self-excusing) and I get angry when someone thwarts me, I become stubborn when I am thwarted, I have a grudge against those who consciously do wrong against me. The factors of interference and self depreciation are clearly indicated. The social positive response 14 may have self reference—I pity unfortunates. The social positive response 5 isn't altogether a positive one—I hate nobody in particular. It suggests that the hate response is there but that it has not been focalized upon any particular individual. The ego positive responses shown in 3, 11, 8, 10 refer in part to potential achievements—I brag about my marks when they're good, I feel proud when I do something worth while, I make believe I finished school—and in part to a passive recipiency—I feel happy when I am pleased. The ego positive responses seem to link up with the ego negative attitude in a sort of compensatory capacity. A fear-worried-unhappy state is indicated by the responses 7, 9, 16, 20—I worry over little things, I am afraid when something goes wrong, I feel sorry when I see tragedy, I wish for happiness. A love attachment is seen in responses 1 and 4—my hero is Oscar, I love Oscar. An intellectual interest is seen in responses 17—I like to read.

Comment: A very strongly organized ego negative and social negative attitude, a fear-worried-unhappy state and a love attachment, a possible tension-relieving activity in reading.

Case J. The perseveration feature is here significantly illustrated. There occurs a perseveration with respect to being

taken advantage of in responses 1 and 5—my hero is one who never takes advantage over anybody, I hate to see a person taken advantage of. Perseveration with respect to the sex motive appears in responses 10, 13, 12, and 18—I brag about my free and easy manner with men but my friends know me, I become stubborn when I feel that I have been too easy going, I have a grudge against the male sex in general and against one in particular, I become disgusted with conceited fool males. Perseveration with respect to unhappiness appears in responses 8, 9, 15, 20—I make believe I'm happy when I feel most like bawling, I feel sorry when moments of happiness are over, I feel ashamed when one discovers that underneath my assumed air of gaiety there is an air of misery sometimes, I wish for unalloyed happiness for at least a year and then it doesn't matter what happens to me. It may well be that the three perseverations are related. Positive social tendencies are indicated in responses 3 and 11—I feel happy when people I like are with me, I feel proud when someone I like achieves something. Ego negative indications may be seen in responses 6, 14—I feel hurt when I am accused of some petty thought which I never entertain, I pity people with inferiority complexes. A source of annoyance is indicated in response 2—I get angry when I see people wrangle constantly. Response 4—I love people who are agreeably sarcastic—strikes at a similar situation but adds an intellectual element and an agreeable element. A fear is seen in respose 16—I am afraid when I see lightning. A social-intellectual interest is indicated in response 17—I like to go to the theatre. Response 19 is suggestive of the imaginative liar—I tell lies when I feel that my imagination is stagnating.

Comment: Perseveration feature is significantly illustrated—a possible tie up of 3 perseverations—being taken advantage of, sex motive, and unhappiness. Ego negative attitude, unwholesome environmental situation of quarrelsomeness, possible imaginative liar.

A VALIDATION OF THE TEST

1. Comparison of emotional insight test with biographic sketches

Thus far there have been presented the behavor of individuals on the test and the possibilities for interpretation. To obtain

some indication as to how responses on the Emotional Insight Test are related to life situations a group of subjects were asked to write character sketches of themselves in addition to taking the Emotional Insight Test. A few samples are here given. Samples were taken at random from different parts of the distribution.

CASE 1. YOUNG WOMAN, MARRIED, AGE 23, WOODWORTH PERSONAL DATA 24 UNFAVORABLE RESPONSES

Tendler Emotional Insight Test

1. My hero is—success
2. I get angry when—things do not go right
3. I feel happy when—everything goes right
4. I love—people, animals, beauty
5. I hate—noise, quarreling, confusion
6. I feel hurt—when people cross me
7. I worry over—everything
8. I make believe—I am very successful and everyone likes me
9. I feel sorry when—I see people ill, poor, unhappy
10. I brag about—my success
11. I feel proud when—I do something nice and am praised
12. I have a grudge against—slowness
13. I become stubborn when—I do not get my own way
14. I pity—people in trouble
15. I feel ashamed when—I do wrong things
16. I am afraid when—I am in the dark
17. I like to—work and do things well
18. I become disgusted with—everything
19. I tell lies when—I think it is best
20. I wish for—success, more happiness

Subject's character sketch

It seems now as if I am everything wrong. I am easily hurt and discouraged. I am tired all of the time. When I stop to think it over, nothing is really wrong but I get so worked up and nervous and going so fast, I do not take time or have the patience to think about it. My husband is very slow, so I have all of the responsibility of a big house, maid, dog and his business, as well as my own. My husband's being slow and not very successful makes me think only in terms of success and makes me very quick. Nothing ever bothers him. He is extremely calm and happy-go-lucky and instead of calming me it does the opposite. I am now wondering what the end will be. I am always very cheerful

and my friends envy me I know. I am not self-conscious and I do everything I am interested in well, such as keeping house, my business, playing bridge, driving, etc. I am praised all the time, but I am mentally in a turmoil all of the time. I can never relax.

Comments: There is considerable agreement between the subject's character sketch and the Emotional Insight Test responses. The complaint of things going wrong, the obsession on success, the difficulties with the husband, his slowness and failure, the unhappiness, the tension, the satisfaction with herself, the self-assertiveness are brought out in both reports.

CASE 2. YOUNG WOMAN, AGE 21, WOODWORTH PERSONAL DATA SCORE 11 UNFAVORABLE RESPONSES

Tendler Emotional Insight Test

1. My hero is—an intelligent, attractive man
2. I get angry when—I'm laughed at
3. I feel happy when—I get the things I want
4. I love—to be good
5. I hate—no one
6. I feel hurt when—someone slights me
7. I worry over—my future.
8. I make believe—I am very wealthy
9. I feel sorry when—I see someone hurt
10. I brag about—nothing
11. I feel proud when—I am commended
12. I have a grudge against—no one
13. I become stubborn when—I'm laughed at
14. I pity—old age
15. I feel ashamed when—after I do something wrong
16. I am afraid when—I am alone
17. I like to—play bridge
18. I become disgusted with—people who tell lies
19. I tell lies when—they are small
20. I wish for—happiness

Subject's character sketch

I am fairly intelligent, showing no signs of abnormality in either a positive or negative way. I seem to be one of a crowd seldom if ever taking the initiative in any movement but am willing to cooperate with any started by someone else.

I am extremely tall for a girl and as a result feel conscious of it. Many times it hinders me in the way of meeting people. I am afraid that I shall be subject to embarrassment. This really is only when I am to meet members of the opposite sex.

I am very affectionate and show my likes.

Comment: Both the character sketch and the Emotional Insight Test responses show the presence of an ego-negative attitude. On the Emotional Insight Test, responses 2, 6, 10, 11, 13, 15 bear upon this point. Features bearing on this point in the character sketch are the lack of initiative, the self-consciousness and embarassment over being tall, the readiness to follow others. The Emotional Insight Test directs attention to additional features: responses 3 and 8 suggest economic difficulties; response 7, worry over the future, and 14, I pity old age may be linked up with the economic or self-depreciative motive, or possibly with the love motive seen in response 1 and stated in the character sketch in the embarassment which hinders her in meeting individuals of the opposite sex. Indications of a social positive interest are seen in responses 5, 9, 12, 17. The Emotional Insight Test in this case shows clearly the diagnostic value of the test in contrast with the subject's biographical statements.

CASE 3. YOUNG WOMAN, AGE 22, WOODWORTH PERSONAL DATA UNFAVORABLE RESPONSES 6

Tendler Emotional Insight Test

1. My hero is—a clever person
2. I get angry when—disappointed
3. I feel happy when—having good time
4. I love—traveling
5. I hate—visiting Doctors, Dentist
6. I feel hurt when—offended
7. I worry over—studies
8. I make believe—I—
9. I feel sorry when—I do wrong
10. I brag about—great achievements
11. I feel proud when—accomplished something great
12. I have a grudge against—conceited people
13. I become stubborn when—things don't go my way
14. I pity—those inflicted with disease
15. I feel ashamed when—I do something wrong
16. I am afraid when—in danger
17. I like to—indulge in sports
18. I become disgusted with—unpleasant sights
19. I tell lies when—I have to defend myself
20. I wish for—great achievements

Subject's character sketch

I rate myself as an average person. I have an anxiety to accomplish many things but so far have accomplished little, am energetic but not

always efficient. I worry quite a bit about my work which hinders my success. I try to conceal my sentiments when I know my opinion would not be favorable to people. I am very anxious to help people. If I could afford it, I would do charity work. I am rather stubborn but try hard to remedy the handicap.

Comment: The achievement motive stressed in the subject's character sketch appears on the Emotional Insight Test in responses 10, 11, 20; the stubbornness feature may be seen in responses 2 and 13; the worry over work appears in response 7; the desire to help people in response 14. In addition the Emotional Insight Test directs attention to an ego-negative attitude in responses 6 and 12, 9, 15 and 19. An aversion to sickness is suggested in responses 5, 14, 18; an interest in travel, sports, good times, in responses 4, 17, 3.

The possibilities of validation by comparison of the Emotional Insight Test with autobiographical sketches are clearly indicated. Features of traits given spontaneously and undirected in the subject's character sketch appear in the Emotional Insight Test. In addition factors not given in the character sketch appear in the Emotional Insight Test.

2. *Validation by comparison of Emotional Insight Test with Woodworth Personal Data Blank*

A quantitative approach to establishing validity and significance may be had through the comparison of two groups—an emotionally favorably organized group and an emotionally unfavorably organized group.

For this purpose all the individuals of six classes in a first course in Educational Psychology given at two summer sessions at Hunter College were the subjects for the experiment. In all there were 190 young women. Their ages ranged from 17 to 27, with the mode at 19; 76 per cent of the cases fell at ages 18, 19, 20.

The subjects were given the Otis SA Higher Examination, Form A, to determine intellectual status. To determine emotional stability status Woodworth's Personal Data blank, containing 70 questions, was given. The Emotional Insight Test, here reported, was given to the whole group.

Using Woodworth's Personal Data Blank as a criterion the group was divided into four quartiles, the first and last being

chosen as representing favorably and unfavorably emotionally organized groups. Comparison of these two groups in intellectual and Emotional Insight Test behavior was undertaken. The scoring of the Emotional Insight Test responses was in terms of these two contrasting groups. Responses given by individuals of the first quartile, the favorably emotionally organized group, hereafter referred to as the good group, were given a value of 1; responses given by the fourth quartile, the unfavorably emotionally organized group, hereafter referred to as the poor group, were given a value of 3. The final value of each response was the average of the ones and threes it had received. Each response on a test paper was then scored and the total for the 20 questions on each paper was found. To obviate dealing with

TABLE 2

	AVERAGE OTIS	S D	AVERAGE WOODWORTH PERSONAL DATA	S D	AVERAGE TENDLER E I TEST	S D
Poor group	53.7	±11.01	23.38	±3.89	504.57	±30.44
Good group	54.3	±9.86	4.62	±2.13	295.13	±23.26
Differences	0.6	±2.17	18.76	±0.637	209.44	±5.67
Diff/SD	0.36		29.4		36.9	

decimals the score was multiplied by 10. A high score would therefore be indicative of poor response, a low score of good response.

In table 2 are given the quantitative features for the poor and good groups: averages, variability and reliability of differences for the Otis, Woodworth, and the Emotional Insight Test.

Table 2 indicates that the poor and good group are not significantly different in intelligence, in fact are closely similar, but are significantly differentiated by emotional tests. Both the Woodworth Personal Data Blank and the Tendler Emotional Insight Test show significant differences, the latter being somewhat more discriminative. We have here an indication on the quantitative side of the validity and significance of the Emotional Insight Test.

SUMMARY

1. Criticism of verbal emotional tests may be directed towards an evaluation of their usefulness in terms of whether they provoke emotional or thought discriminative responses.

2. Criteria are suggested for emotional tests in terms of direct evocation of emotional responses, non-discriminative reaction, and free response.

3. A new test based on the completion test technique and directed towards provoking Emotional Insight is presented.

4. Characteristic features of the behavior on the Emotional Insight Test are educed from a sample of cases.

5. The possibilities for individual analysis in terms of the features educed are indicated.

6. Indications for a validation (qualitative) of the test are given from a comparison of Emotional Insight Test responses and character sketches of themselves written by the subjects.

7. Indications of a quantitative validation are given from a comparison of a favorable and unfavorable group based on Woodworth's Personal Data Blank.

DISCUSSION

The particular instrument and method used opens up research possibilities in the study of emotional and personality development. Comparative study of children's and adult responses is one direction in which data are now being gathered. Another is group differences. In this direction data are being gathered from a group of unmarried mothers, neurotic subjects, different age and sex groups.

Clinically the instrument has been found to be of value as a device for eliciting attitudes, trends, and significant clues to be followed up by further questioning. The conversational approach follows naturally and coöperatively with the individual explaining more fully what he meant by his responses. A valuable feature of the test is that it does not suggest a pathological turn so that it does not evoke an antagonistic attitude on the part of the subject. Theoretically, this feature is of value in that it does constitute an approach to emotional behavior in its normal phases. The same advantages would seem to hold for its use in counselling and personnel work.

VARIABILITY AND CENTRAL TENDENCY IN AESTHETIC JUDGMENTS

NATHAN ISRAELI

Columbia University

Varieties of aesthetic experience are subject to the law of central tendency. Individual differences tend to disappear with increasing age as shown in a comparison of the variability of aesthetic preferences of different age-level groups. Data on variability and central tendency in aesthetic judgments of colored triangles and painting reproductions in colors will be reported here. Such data will be fundamental to a genetic study of aesthetics.

GOLDEN-SECTION EXPERIMENT (TRIANGLES IN COLORS)

This experiment was limited to triangles. A complete set of five different color series (black, red, green, yellow, blue) for the same twelve triangles was provided each subject. The figures were printed from a cut made from Thorndike's *Education-A* manual and are of the same dimensions as those in his 1917 investigation.[1] In the present series the original numerical designations (41–52) were changed (1–12). A standard procedure was followed. The instructions printed on the data sheets follow: "Opposite each *position-number* below (12 in a series) write the number of a *figure* according to how pleasing the figure is to you. Write the number of the most pleasing figure opposite position-number 1, and so on down to the number of the least pleasing opposite position-number 12." The time was controlled. Ninety seconds were given for a figure-series with 30 seconds rest between consecutive figure-

[1] Edward L. Thorndike, Individual differences in judgments of the beauty of simple forms. Psy. Rev., *24*, 1917, 417–153.

series. One hundred sixty-two subjects were tested. They included 47 graduate students[2] (taking a mental-tests course) with an age average of 32 years—4 months (an age range of 19–55 years) and three freshmen high-school groups[3] with age average of 14 years 4 months. For the graduate students the order in the experiment of the figure series was black, red, green, yellow, blue; that order was reversed in Group-I (50 subjects) of the high-school students, held for Group-II (35 subjects), but was changed in Group-III (30 subjects) to blue, red, yellow, black, green.

CENTRAL TENDENCY (GOLDEN-SECTION EXPERIMENT)

In this statistical treatment only the first five figures (1–5) in the first four positions (order-of-merit) will be examined. The A.P.A.'s (average percentage agreements) for the first four figures in these positions would be equivalent, the A.P.A.'s equaling 33.3, if no factors other than chance or the same-figure-same-position tendency are operative. This tendency is present in the last ten places; figures 3–4 are the only exceptions and precede 1–2 both in this experiment and in Thorndike's. That the A.P.A.'s are not equivalent is significant. Only the first five figures (41–45) were assigned by more than 33.3 per cent of Thorndike's 250 college juniors to the first four positions. Figure 1 shows the definite inversion of the parallelism of the groups after the first five or six figures. Furthermore, more than 50 per cent of the juniors assigned the first six triangles to positions 1–6. The validity of the method cannot be gainsaid.

The criterion of 33.3 in the A.P.A.'s is no arbitrary one. It is a measure of zero-aesthetic judgment. The operation of chance factors alone would cause an equal selection of all the figures for all the positions by the same number of judges. Assuming chance selection only, since there are twelve positions in the present order-of-merit one-twelfth of the members of a group

[2] Through the courtesy of Professor R. Pintner, Teachers College, Columbia University.

[3] Through the courtesy of G. C. Hanson, Boys High School—Waverly Annex (Brooklyn, N. Y.).

would select figure 1 for position 1; another twelfth, for position 2; another twelfth, for position 3; and a fourth twelfth would select figure 1 for position 4. Altogether four-twelfths of the group would select figure 1 for the first four positions. And that is the origin of the 33.3 A.P.A. criterion. In table 1 the

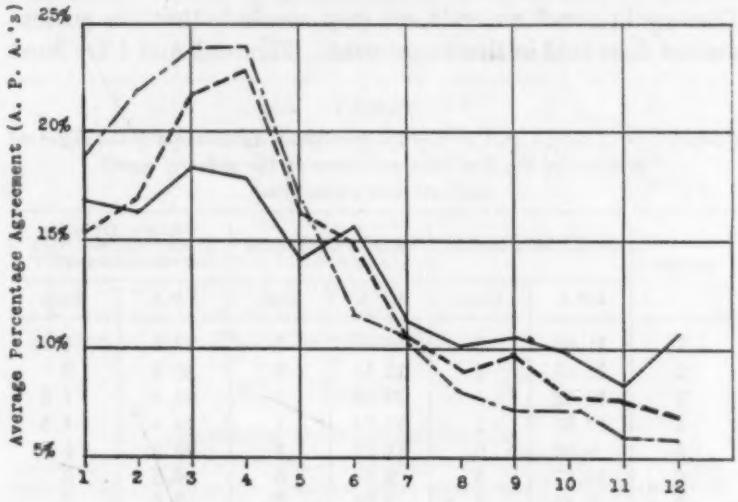

FIG. 1. PARALLELISM IN THE GOLDEN-SECTION (TRIANGLES IN COLORS) EXPERIMENT

Showing the differentiated reactions of different age-level groups. The dot-and-dash line represents the judgments of the college-juniors (computed from data of Thorndike, 1917); the dash line, graduate-students (age range: 19–55 years); the unbroken line, high-school subjects. 412 subjects.

A.P.A.'s for the first four positions can be converted into total P.A.'s through multiplication by 4. The total P.A.'s exceed 33.3 in the case of the first five figures. Figures 3 and 4, the golden-section triangles, have the highest total percentage agreement for the first four positions.

That this method is a good criterion of any golden-section appeal must needs be supplemented by other controlled experi-

ments in which the separate triangles are evaluated individually or with twelve triangles omitting the golden-section ones (3–4 here). In the latter case if the figures 3–4 of such an experiment attracted numerically as high a proportion of the subjects as the golden-section figures 3–4 do now the conception of the golden-section reaction by law of parsimony would be ruled out. Disregarding such controls, we may conclude that the golden-section does hold in this experiment. Figures 3 and 4 are most

TABLE 1

Comparison of the (A.P.A's) average percentage agreements of the different groups for the first four positions in the order-of-merit with all colors combined

FIGURE	HIGH-SCHOOL GROUPS		GRADUATE STUDENTS		COLLEGE STUDENTS (COMPUTED FROM DATA OF THORNDIKE, 1917)	
	A.P.A.	Rank	A.P.A.	Rank	A.P.A.	Rank
1	11.95	3	10.38	5	13.8	4
2	11.45	4	12.11	3	16.8	3
3	13.42	1	16.36	2	18.8	1.5
4	12.95	2	17.74	1	18.9	1.5
5	9.10	6	11.25	4	10.2	5
6	10.67	5	9.88	6	6.6	6
7	6.42	7	5.74	7	5.6	7
8	5.13	9	4.04	9	3.1	8
9	5.53	8	4.88	8	2.2	9.5
10	4.80	11	2.87	10	2.2	9.5
11	3.35	12	2.34	12	1.1	11
12	5.26	10	1.80	11	.9	12

pleasing to all the subjects. The data on the college students presented here (cf. table 1) were compiled from Thorndike's order-of-merit table.[4] There is close agreement with the results of this experiment.

The subjects behaved very much alike. Analysis of the results indicates quantitative differentiation due to the influence of individual differences presumably unlike age-levels. A comparison of the average percentage agreements of the different

[4] Op. cit., p. 151.

groups in the first four positions in the order-of-merit with all colors combined in table 1 shows this qualitative similarity and

TABLE 2

Reliability of the inter-differences (in the comparison of the A.P.A's of the various groups)

FIGURE	HIGH-SCHOOL GROUPS* VS. GRADUATE STUDENTS†	COLLEGE STUDENTS‡ VS. GRADUATE STUDENTS	COLLEGE STUDENTS§ VS. HIGH-SCHOOL GROUPS
1	.26	.79	.50
2	.28	.96	1.19
3	1.73	.32	1.25
4	2.63	.25	1.38
5	1.58	.54	.70

* N = 60 and based on 115 subjects.
† N = 20 and based on 47 subjects.
‡ N = 4 and based on 250 subjects.
§ Computed from Thorndike's data, 1917.

TABLE 3

Reliability of the intra-differences $\frac{D}{\sigma_D}$

FIGURES COMPARED	HIGH-SCHOOL GROUPS*	GRADUATE STUDENTS†	COLLEGE STUDENTS‡§
1-2	.35	.51	.53
2-3	1.67	1.61	.33
3-4	.40	.58	
4-5	3.38	2.86	1.98
3-1	1.13	2.76	.65
3-5	4.45	2.46	1.92
4-1	.70	3.10	.93
4-2	1.13	2.00	.35

* N = 60 and based on 115 subjects.
† N = 20 and based on 47 subjects.
‡ N = 4 and based on 250 subjects.
§ Computed from Thorndike's data, 1917.

quantitative stratification of the reactions to the triangles. The high-school groups give minimum P.A.'s and the graduate students lag somewhat behind the college juniors. This

parallelism is graphically shown in figure 1. That these differential P.A.'s are reliable, that the parallelism is significant, is to be concluded from tables 2 and 3 which summarize the reliability of the inter- and intra- differences respectively. The formula [D/s_D] was applied throughout. In determining the reliability for the first places in Thorndike's order-of-merit as N = 4 only the significance of the results may be under-estimated.

The reliability tables should be examined for critical comparisons only. The chief concern here is with the gold n-section triangles 3 and 4 as contrasted with the others and only such critical comparisons are relevant.

Outstanding indeed are the golden-section triangles. Since these triangles (3–4) individually affected all the subjects more than any other triangle, it may be concluded that the subjects gave similar reactions differentiated according to age-levels.

VARIABILITY

Variability of triangle preferences in the various groups is maximum for the younger subjects. Coefficients of variability for figures 1–5 only in the first four positions with all colors combined are grouped in table 4. In the last column is a tabulation of the average variability coefficients according to the divisions. Variability as determined by the rank order of the average variability coefficients is in inverse order to the central tendency determined by the A.P.A.'s. This is confirmed by an analysis of the college juniors' results. They yielded the highest percentage agreement and the lowest variability coefficient. The high-school groups are 119 per cent as variable as the college juniors. The variability of the various groups approaches equality at the golden-section and is there minimum for the high-school and graduate students. Table 4 also includes a high-school group given the same figure-series color order as the graduate students. There is minimum variability in the older groups. The differences between the graduate students and the college juniors are not reliable and the adults may be more or less variable than the college students. A comparison of the A.P.A.'s of the two groups for

the first five figures is found to be only somewhat significant with the differences over the standard errors of the differences not exceeding .79, .96, .32, .25, and .54 respectively. If this is due to the low value of N the college-students would perhaps

TABLE 4

Coefficients of variability (for figures 1–5 only and the first four positions in the order-of-merit with all colors combined)

GROUP	FIGURES					AVERAGE VARIABILITY-COEFFICIENTS FOR FIGURES 1–5
	1	2	3	4	5	
Combined high-school groups	70.36	63.22	41.08	44.32	56.04	55.00
HS-II	47.27	97.77	37.81	49.31	53.89	57.21
Graduate-students	45.29	79.19	42.17	45.99	54.13	53.35
Average variability-coefficients for all groups	54.31	80.06	40.35	46.54	54.69	
Rank	3	1	5	4	2	
*College students**	*51.30*	*52.45*	*45.15*	*44.44*	*27.45*	*44.16*

* Computed from data of Thorndike, 1917.

TABLE 5

Colors of the most pleasing figures

HIGH-SCHOOL GROUPS				GRADUATE STUDENTS			
P.A.'s*	Color	P.A.'s†	Color	P.A.'s*	Color	P.A.'s†	Color
115.3	Blue	182.0	Red	125.0	Black	226.0	Red
113.5	Red	176.0	Blue	89.0	Red	178.0	Black
99.0	Green	172.0	Black	88.0	Green	170.0	Green
94.0	Yellow	158.0	Yellow	81.0	Blue	108.0	Blue
90.0	Black	140.0	Green	60.0	Yellow	106.0	Yellow

* Most pleasing color per figure.
† Most pleasing figure per color.

show less agreement and more variability than the graduates. This is conjectural. Nevertheless both groups are less variable (some doubt in the case of the graduate students) than the high-school freshmen and yield reliably higher P.A.'s.

The conclusion is therefore reached that in aesthetic preferences of certain triangles there is an inverse relationship between central tendency and variability with the former minimum for the high-school, high for the graduate students and maximum for college juniors (reliability of the differences between the latter two slight only).

COLOR SHIFT EFFECTS

That the golden section is selected regardless of color is concluded from the data in table 1. Shift effects in the selections are surmisable from table 5. Yellow is the least pleasing color for all groups and probably cuts against the golden-section form whereas red considerably affects the results and is the most pleasing color facilitating perhaps the golden-section appeal. This may be seen in an inspection of the data in the columns entitled the most pleasing figure per color where red is first. In the columns entitled the most pleasing color per figure it is second. The latter columns were obtained by summing up the P.A.'s for the groups specified of the most pleasing colors per figure in the first four positions in the order-of-merit; and the former, by totalling the P.A.'s of the most pleasing figures per color.

Variability and Central Tendency in Judgments of Painting Reproductions

Application of the analytic method used above to the present writer's previous study of affective reactions to painting reproductions in colors[5] puts to the test the conclusions on diminished variability and accelerated central tendency arrived at in the golden-section experiment. A striking parallelism between the college, high-school, and art divisions is clearly apparent in table 6. The averages for the divisions of the combined A.P.A.'s of 10 or over for the twenty pictures and twenty-three adjectives are 25.45 (college), 20.68 (high-school) and 19.55 (art). The differences in these averages of the combined

[5] Jr. Appl. Psy., *17*, 1928, 125–139.

TABLE 6

Combined average P.A.'s of 10 or over for the pictures according to the division

PICTURE	COLLEGE DIVISION	HIGH-SCHOOL DIVISION	ART DIVISION
1	34.4	30.0	23.9
2	16.4	16.6	13.1
3	25.9	23.4	19.4
4	38.7	29.6	24.9
5	31.4	26.2	17.6
6	34.8	21.9	21.0
7	17.0	14.8	11.8
8	24.6	15.4	16.6
9	16.6	10.9	13.8
10	27.1	21.7	21.3
11	22.4	20.7	17.4
12	23.2	20.8	20.0
13	22.3	16.9	21.9
14	29.2	23.8	25.2
15	29.5	32.1	24.7
16	18.8	18.6	22.8
17	29.8	23.5	16.3
18	21.8	10.8	20.7
19	27.4	20.1	25.5
20	17.8	15.9	13.2
Av.....	25.45	20.68	19.55
σ.....	6.303	5.802	4.262
V.....	24.76	28.05	21.8

TABLE 7

Reliability of differences in judgments of pictures

GROUPS COMPARED	$\frac{D}{\sigma_D}$
College-high school.....	2.49
College-art.....	3.47
High school-art.....	.70

A.P.A.'s (of 10 or over) are highly significant in comparing the college and high-school divisions (table 7). The chances of a true difference greater than zero between the average of the combined A.P.A.'s of the high-school and art divisions are however only 77 in 100.

TABLE 8

How the subjects behave alike with parallelism of the divisions effective

PICTURE	ADJECTIVE SERIES	PERCENTAGE AGREEMENT			A.P.A.
		College division	High-school division	Art division	
2	(7) tumult	19.7	17.6	11.7	16.3
	(2) depression	14.6	15.4	18.4	16.1
	(5) excitement	13.1	15.8	6.2	11.7
	(1) calmness	8.6	20.9	5.0	11.5
7	(7) tumult	19.2	24.9	15.9	20.0
	(5) excitement	17.6	18.5	14.6	16.9
	(2) depression	16.0	8.8	8.6	11.1
	(1) calmness	9.0	14.8	6.1	10.0
9	(7) tumult	24.0	19.1	21.1	21.4
	(13) sprightly	17.0	4.2	22.8	14.7
	(1) calmness	16.0	17.8	5.0	12.7
	(5) excitement	10.6	9.6	11.1	10.4
11	(1) calmness	36.0	37.0	35.5	36.2
15	(1) calmness	41.4	42.7	18.6	34.2
	(13) sprightly	24.3	18.1	33.9	25.4
20	(7) tumult	9.4	7.7	22.8	13.3
	(5) excitement	8.7	6.2	15.3	10.1
	(1) calmness	5.3	21 1	3.7	10.0

Conclusions 2–3 in the former paper[6] are therefore correct. A horizontal comparison in table 8 of the A.P.A.'s of the art, college, and high-school divisions demonstrates the clear-cut parallelism and synchronous deviations of the emotional reac-

[6] *Ibid.*, p. 138.

tions on the part of the different groups. This table was derived from table 1 in the former report and its first three columns—

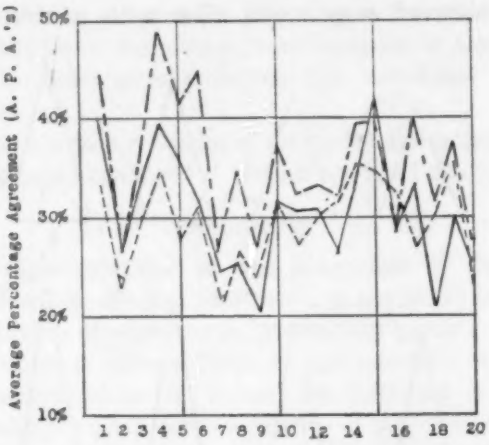

FIG. 2. PARALLELISM IN THE COLORED PICTURES EXPERIMENT

Showing that the 400 subjects behaved very much alike with differentiated reactions of the various divisions. The dot-and-dash line represents the judgments of college students; the dash-line, graduate-art students; the unbroken line, high-school subjects.

TABLE 9

Relativity of Variability (comparison of variability of college and high-school students in aesthetic judgments of pictures and triangles)

	PICTURES*	TRIANGLES†
College students	24.76	44.16
High-school students	28.05	55.00

* Derived from table 6.

† Derived from table 4.

from the original unpublished table of 1380 P.A.'s for the three divisions[7] (from which the table for all the subjects was de-

[7] M. A. Thesis, 1927, Columbia University Library.

rived[8])—supplement the final column of excerpts made by a critic[9] from table 2 of the previous report. All the tables in this section of the paper and figure 2 demonstrate that the subjects behaved very much alike with a stratification of

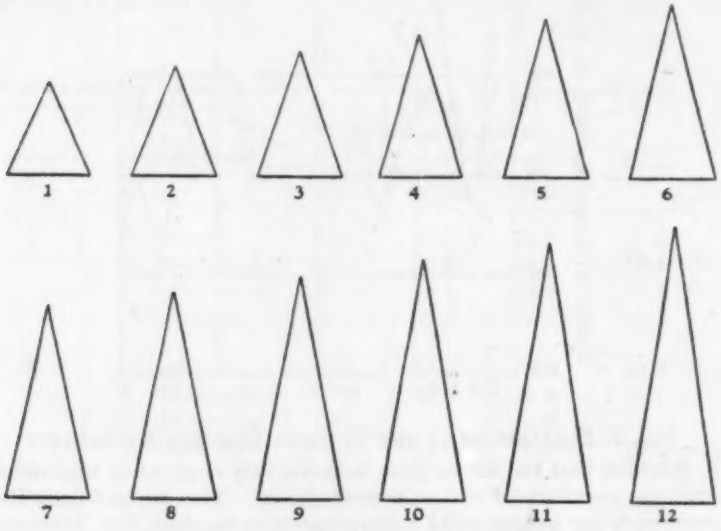

FIG. 3

reactions according to the different divisions. In the light of the data here presented it is perfectly evident that the criticism of the former paper's conclusion that the 400 subjects behave very much alike "seems to be an overstatement"[10] is without

[8] Table 1 in the 1928 report requires the following corrections:

Adjective	*Picture*	*A. P. A.*
3	5	32.2
10	8	13.7
12	10	8.7
8	19	.9

Table 2 corrections to be made from table 1.

[9] Albert R. Chandler, Recent experiments on visual aesthetics, Psy. Bul., *25*, 1928, p. 730.

[10] *Ibid.*, p. 729.

basis. Although differential emotional responses were exhibited, there was similarity of response all the way through.

Coefficients of variability computed for the combined A.P.A.'s of 10 or over for all the pictures according to the divisions show distinctly that there is a concomitant variation of age and variability. The older groups showed less variability than the younger ones.

Table 9 points to a *relativity of variability* in aesthetic differential percentage agreement of various age-level groups.

CONCLUSIONS

Graphic representations of the judgments of the various groups confirm the observations made in the previous sections about parallelism of reactions in different age-levels.

In preferences of colored triangles and affective reactions to painting reproductions in colors the following conclusions were reached: (1) qualitative similarity of reactions is the rule for all subjects ("the subjects behave very much alike"); (2) quantitative differentiation is found in the central tendency and variability (with inverse relationship between the two) of different age-level groups. The law of central tendency holds for aesthetic judgments: with increasing age there is less variability and more agreement.[11]

[11] Compare with work of Marion P. Jenkins (Lowell) referred to in the present writer's previous paper.

SCORES IN THE PRESSEY X-O TESTS OF EMOTIONS AS INFLUENCED BY COURSES IN PSYCHOLOGY

JOHN A. McGEOCH
University of Arkansas
AND
MARION E. BUNCH
Washington University

The purpose of this investigation is to study the influence of the teaching of certain topics in psychology upon scores on the Pressy X-O Tests for Investigating the Emotions. It was found in a previous study[1] that the reliability of these tests decreases considerably with increasing time interval, and that the qualitative constancy of response likewise decreases. The 90-day group used in that investigation had been studying such topics as emotion, personality, personality maladjustments, and mental hygiene, and their test scores may have been influenced thereby. It was suggested that the progressive change in the constancy of the test scores might be explainable in terms of changed affective organization resulting from progressive alteration of the stimulating conditions to which the subjects were exposed during the interval. The study of the topics mentioned was presumably a significant part of these stimulating conditions.

In the present investigation an attempt has been made to discover the influence of the teaching of such topics upon Pressey X-O scores. The Pressey X-O Tests, Form A, were given to four sections of sophomores in the second semester of general psychology. The 20 lectures of the succeeding 45 days were, with a control group of two sections numbering 66 students,

[1] McGeoch, John A., and Whitely, Paul L., The Reliability of the Pressey X-O Tests for Investigating the Emotions, Ped. Sem. and J. Genet. Psychol., 1927, 34, 255–270.

devoted to a discussion of topics in business psychology, such as the conditions of work, personnel psychology, advertising and selling. To an experimental group of two sections, numbering 91 students, were given 20 lectures on personality, several forms of personality maladjustment, a brief survey of the psychoanalytic principles of Freud, Adler, and Jung, particularly as they bear on the organization and motivation of the individual, and finally mental hygiene. A liberal use was made of case histories and collateral reading.[2] At the end of 45 days, during which no mention had been made of the Pressey tests, they were given again. *A priori* it would be expected that a study of the topics just enumerated would bear much more closely on the affective organization of students than would the study of business psychology. The problem, then, is to compare the Pressey scores of the two groups at the end of the 45 day period, to determine whether any significant differences appear between them. If such differences are found, they may reasonably be ascribed to the differences in subject matter to which the groups were exposed during the interval.

The different conditions to which the groups were subjected are admittedly rough and impossible of quantitative statement. They depend greatly upon the material studied and the skill of the teacher. Particularly is this true of the group presented with a discussion of personality and kindred topics. It would be desirable to specify in detail the material with which this group was presented, but this to be useful would require a too comprehensive syllabus to permit statement here. Even such a statement would be of uncertain significance apart from a knowledge of the specific conditions prevailing in the course, such as the motivation of the students, the methods of the lecturer, and the thoroughness with which the collateral reading was done. The conditions studied are, however, widely divergent. Business psychology as presented has little in common with personality, its maladjustment and hygiene. It is,

[2] The sections of the control group were taught by Professor Bunch; those of the experimental group by Professor McGeoch.

therefore, a valid question whether two different aspects of psychology presented to different groups of students over identical periods of time will be accompanied by differences in Pressey scores. We may fairly assume that other conditions would be roughly equal in groups of a considerable size from the same college class. If differential results are obtained, it will be indicated that Pressey scores and hence, presumably, affective organization can be affected by the types of knowledge here presented.

SCORING AND TREATMENT OF RESULTS

Affectivity and idiosyncrasy scores have been obtained for each test, and the words crossed out in Tests 1 and 4 have been analyzed according to the classification schemes in which they are arranged. The means and standard deviations of these scores have been found, together with the Pearson coefficient of variation and the relative variability of the two givings, obtained by dividing the coefficient of variation for the second giving by that for the first giving. Pearson product-moment coefficients of correlation have also been computed between the two givings of the tests for each group. The qualitative change or stability for each group has been studied by finding the numbers of crossed out and circled words not common to both givings of the tests.

AFFECTIVITY SCORES

The affectivity scores, their standard deviations, Pearson coefficients of variation, and relative variabilities are given in table 1.

In every case the means for both groups show an increase on the second giving of the tests. The significant fact is, however, that with the control group the coefficient of variation of every test decreases from the first to the second giving, while with the experimental group the coefficient of variation of Tests 1 to 3 increases and that of Test 4 falls slightly. For every test the relative variability of the second giving in comparison with the first is greater for the experimental than for the control group. The relative variabilities average 93.10 for the control group and

103.92 for the experimental group. Both these means and the consistency of the direction of the differences point to an increased variability resulting from the experimental conditions applied.

TABLE 1

Affectivity scores

		CONTROL GROUP N—66		EXPERIMENTAL GROUP N—91	
		0	45 days	0	45 days
Test 1	Mean	38.51	42.00	40.07	44.37
	S.D.	16.50	15.45	13.80	16.10
	V	42.84	36.78	34.48	36.28
	Rel. V	85.86		105.22	
Test 2	Mean	49.80	52.07	49.43	51.87
	S.D.	18.00	18.20	14.30	15.70
	V	36.14	34.95	28.92	30.26
	Rel. V	96.70		104.63	
Test 3	Mean	67.98	70.25	66.58	69.99
	S.D.	14.30	14.30	16.50	18.60
	V	21.03	20.35	24.78	26.57
	Rel. V	96.76		107.22	
Test 4	Mean	42.46	44.52	45.98	46.63
	S.D.	21.00	20.50	19.90	19.90
	V	49.45	46.04	43.27	42.67
	Rel. V	93.10		98.61	

IDIOSYNCRASY SCORES

The results of a similar treatment of the idiosyncrasy scores are given in table 2.

The mean idiosyncrasy scores of all of the control group tests increase with the time interval, while those of the experimental group vary in direction. The mean of Test 1 increases, and those of Tests 2 to 4 either decrease or remain constant. The variability of the second testing relative to the first is higher for the experimental than for the control group in three out of

the four cases, and the mean of the relative variability scores of the former is 98.48 and that of the latter is 106.38. Again the experimental conditions have, apparently, been effective in generally increasing the variability of the scores of the experimental group and in producing a slight tendency for the direction of change of the means to reverse that of the control group.

TABLE 2

Idiosyncrasy scores

		CONTROL GROUP		EXPERIMENTAL GROUP	
		0	45 days	0	45 days
Test 1	Mean	10.05	10.50	10.63	10.86
	S.D.	2.56	2.68	3.40	3.02
	V	25.47	25.52	31.98	28.26
	Rel. V	100.19		88.36	
Test 2	Mean	8.96	9.44	10.15	9.87
	S.D.	2.40	2.26	2.82	2.92
	V	26.78	23.94	27.78	29.58
	Rel. V	89.54		106.48	
Test 3	Mean	13.08	13.81	14.12	14.12
	S.D.	2.56	2.66	2.84	2.96
	V	19.57	19.26	20.11	20.96
	Rel. V	98.41		104.22	
Test 4	Mean	15.87	15.99	16.22	16.09
	S.D.	2.44	2.60	2.28	2.86
	V	15.37	16.26	14.05	17.77
	Rel. V	105.78		126.47	

CLASSIFICATION SCHEMES

The tendencies of the classification schemes of Tests 1 and 4 mirror so closely the results shown in table 1 that it is scarcely necessary to give them here. In Test 1 the differences between the variabilities of the two groups is greatest for sex words, with fear, disgust, and self-feeling words following in the order named. In Test 4, all but melancholic words show a decrease in vari-

ability after the time interval with the control group, while with the experimental group all but shut-in words show an increased variability.

COEFFICIENTS OF RELIABILITY

It is significant to inquire further whether the experimental conditions have affected the relationships between the two

TABLE 3

Reliability coefficients—affectivity scores

	CONTROL GROUP		EXPERIMENTAL GROUP		D	$P.E._D$
	r	P.E.	r	P.E.		
Test 1	.689	.043	.572	.047	.117	.063
Test 2	.857	.021	.683	.038	.174	.043
Test 3	.692	.043	.744	.032	.052	.053
Test 4	.800	.030	.725	.034	.075	.045

TABLE 4

Reliability coefficients—idiosyncrasy scores

	CONTROL GROUP		EXPERIMENTAL GROUP		D	$P.E._D$
	r	P.E.	r	P.E.		
Test 1	.208	.080	.704	.036	.496	.087
Test 2	.375	.071	.478	.054	.103	.088
Test 3	.427	.068	.628	.043	.201	.080
Test 4	.361	.072	.415	.059	.054	.093

givings of the tests. The coefficients of correlation between the affectivity scores at the first and second givings for each group (reliability coefficients) are shown in table 3.

The reliability coefficients of the experimental group are lower than those of the control group in three out of the four cases, that for Test 3 only being an exception. The differences between the coefficients of the two groups are, however, unreliable except in the case of Test 2. Nevertheless the fact that three

out of the four coefficients are lower for the experimental group and that the fourth, which contradicts this tendency, is the most unreliable of all, would indicate that the direction of the difference is toward a lowering of reliability by the experimental conditions.

Similar data for the idiosyncrasy scores are shown in table 4.

These reliability coefficients reverse the tendency shown by those for the affectivity scores, the coefficients being in every case lower for the control than for the experimental group. Save for Test 1, however, the differences between the coefficients of the two groups are statistically unreliable. The consistency of their direction, nevertheless, indicates a probable tendency to an increased reliability in the tests of the experimental group. If the reversal of tendency here noted is not a chance reversal, it is necessary to assume that it is the result of the experimental conditions.[3] It is, of course, possible for the affectivity and idiosyncrasy scores to vary independently. Idiosyncrasy scores are computed from the number of times that the word circled by the subject in each line is not the word most commonly circled by Pressey's original group. Further, the circled word in each line is the word to which the subject's affective response is strongest. Presumably the affectivity scores represent the comparatively more superficial affective responses. It is entirely possible that by calling the subjects' attention to their own affective organization the study of the topics described should change considerably the more superficial affective reactions and thus the affectivity score relationships, while not only effecting a much smaller change with a majority of sub-

[3] It should be noted that the general increase in the dispersion of the scores of the experimental group will have operated to increase the correlation. It is significant, therefore, that in spite of this fact the affectivity coefficients of correlation are lower for the experimental than for the control group. Further, the fact that both affectivity and idiosyncrasy scores, in the case of the experimental group, tend to show an increased variability at the second giving, while the coefficients of correlation vary independently of this change in variability, argues for the influence of the experimental conditions upon the correlations over and above the effect upon variability.

jects in the stronger affective responses represented by the circled words but also serving to "set" them more firmly.

WORDS NOT COMMON TO BOTH TESTINGS

The mean numbers of crossed out words not common to both givings of the tests afford a statement of the qualitative constancy or inconstancy of the responses to the materials of the tests. These have been found and the percentages which they are of the mean number of words crossed out at each giving have been computed. These percentages are given in table 5.

TABLE 5

Percentages of words not common

	TEST 1		TEST 2		TEST 3		TEST 4	
	0	45 days	0	45 days	0	45 days	0	45 days
Control	20.9	27.6	17.9	23.0	13.3	15.7	29.1	36.5
Experimental	21.6	30.6	20.5	23.8	14.3	18.5	32.2	32.3

TABLE 6

Numbers of circled words not common

	TEST 1	TEST 2	TEST 3	TEST 4
Control	7.66	7.79	8.82	11.36
Experimental	7.95	8.34	10.07	11.09

There appears a consistent tendency for the tests of the experimental group to contain the larger percentages of words not common to both givings. The second giving of Test 4 is the single exception. The differences are small, but their appearance in seven out of eight cases indicates that the influence of the experimental conditions has been in the direction of a greater differentiation of the responses of the subjects.

The numbers of circled words not common to the two givings are shown in table 6. In this case the number circled at one giving but not at the other is identical for the two givings.

therefore the numbers in this table may be multiplied by two to obtain the total number of circled words not common. Since the total number of circled words is constant at 25 it is useless to give percentages.

The number of circled words not common to both testings is greater for the experimental than for the control group in all save Test 4. The differences are small, but their presence in three out of four cases probably indicates validly the influence of the conditions to which the experimental group had been subjected.

SUMMARY

When the Pressey X-O Tests for Investigating the Emotions are repeated with two groups of subjects after an interval of 45 days, during which one, the control group, had been studying topics in business psychology, and the other, the experimental group, had been studying personality, personality maladjustment, and mental hygiene, the following comparisons between the two testings of each group appear.

1. The variability of the affectivity scores of the second giving relative to that of the first is higher for the experimental group in each of the four tests.

2. The variability of the idiosyncrasy scores of the second giving relative to that of the first is higher for Tests 2, 3, and 4 with the experimental group and for Test 1 with the control group.

3. The reliability coefficients of the affectivity scores on Tests 1, 2, and 4 are lower for the experimental than for the control group.

4. The reliability of the idiosyncrasy scores of the experimental group is, contrary to the tendency of the affectivity scores, higher for each of the four tests than is that of the control group.

5. The tests of the experimental group have a larger percentage of crossed out words not common to both givings, and a larger number of circled words not common, with the exception of Test 4 in both cases.

6. The influence of the teaching conditions prevailing with the experimental group has been, in general, thus, to increase the variability of the affectivity and idiosyncrasy scores, to decrease the reliability of the affectivity scores and to increase that of the idiosyncrasy scores, and to increase the qualitative inconstancy of response indicated by the crossed out and circled words not common to both givings. When it is remembered that the interval studied is but 45 days and that, during this time, the only known differences between the stimuli to which the groups were subject were the teaching conditions previously described, it is surprising that differences as uniform as these should result. The direction of these differences, with but few exceptions, indicates that the teaching of the topics described has a distinct influence on the affective organization of the students as represented in their Pressey scores. This result indirectly supports the supposition, previously stated, that change with time in Pressey scores is, in part at least, the result of changed stimulating conditions. It also points to an important but often unnoticed by-product of teaching; i.e., altered affective organization.

SOME RESULTS SECURED IN A PSYCHO-EDUCATIONAL CLINIC

PAUL A. WITTY
University of Kansas

The Clinic in Educational Psychology at the University of Kansas was established by Professor Raymond A. Schwegler in 1912; it has a two-fold purpose: (1) to train selected graduate students to assemble and interpret psychological and educational data relative to the special needs of school children; (2) to provide educational diagnosis and psychological service for the schools of Lawrence, Kansas, and certain neighboring communities.

In the course of its development the clinic has undergone numerous changes, but the procedure outlined herein has been utilized during the years 1924 to 1928, inclusive, during which time the writer has directed the work.

The students admitted to the clinic training course are graduate students in the University. They enroll for 3 to 5 hours credit for one semester (18 weeks). Seven students only are accepted each year, as facilities limit enrolment. These students are selected in accord with subjective estimates of their likelihood of success in this type of endeavor. The class meets for lecture, laboratory, and demonstration work from 1:30 p.m. to 4:30 p.m., Monday, Wednesday, and Friday of each week. During the first six weeks, three demonstrations are given weekly by the instructor who explains and follows the technique to be employed subsequently by the students. The students learn verbatim the directions to be used in the various examinations. Complete verbal mastery of the technique is required of the student by the end of the sixth week of the semester.

The following twelve weeks are devoted to examination of

cases by the students. Three to five children are referred to the clinic at each meeting. They arrive at 1:30 p.m. and remain until 4:30 p.m. Fatigue is avoided by providing frequent change and periods of rest. Certain cases appear to be unduly fatigued after an hour or two spent in the clinic. These are dismissed; they return two days later to complete the examination.

Each child is sent to the clinic upon request of a teacher in the Lawrence schools or neighboring community. Every child is a problem case sent for one of the following reasons:

a. Failure in one or more subjects.
b. Candidacy for special class.
c. Recent entry to school (problem of classification).
d. Disciplinary difficulties.
e. Nervous instability.
f. Delinquency.
g. Problem of proper classification or special promotion.

Identification data are furnished by the teacher upon each case including:

a. Name—Age—Grade—Nationality of child and of parent.
b. Reason for examination.
c. Scholarship.
d. Home and environmental conditions.

Especial attention is directed toward securing the maximum response from the child. Effort is made to maintain a cheerful, informal atmosphere during the observation. Each child is examined in a separate room, so constructed as to eliminate distracting and conflicting influences. The following types of data are secured during the three hour observation: (*a*) physical, (*b*) mental, (*c*) educational, (*d*) emotional.

a. Physical. Vision and hearing are examined and the condition is recorded. The customary method of scoring, recording, and interpreting the condition of the eyes is employed; i.e., 20/20 vision is normal, etc. Auditory acuity is gauged by the whisper test. Standing height, sitting height, and weight are obtained; these indices are compared with the Baldwin-

Wood standards and diagnosis of the child's condition follows. The throat is examined and recommendations for medical attention prescribed when diseased tonsil or throat is exhibited. During these physical tests, the examiner notes and records any unusual condition of the skin or postural irregularity.

The purpose of this part of the observation is to discern physical limitations which may handicap the child in his school work. Conspicuous weaknesses or anomalies are referred to the school physician for complete diagnosis and treatment.

The examiner is particularly zealous to identify sensory defects which obliterate or distort the imagery of the child. The writer has observed occasional children who displayed a special disability in certain school subjects (notably reading), children whose intelligence was normal or superior but whose vision or hearing was decidedly impaired. In such cases it is essential of course that the sensory defects be identified and corrected before remedial work can be successful.

b. Mental. To gauge the mental ability of the child the Stanford Revision of the Binet-Simon Test of Intelligence is employed. The Porteus Test of Motor Coordination, Prudence, and Forethought is often given to supplement the Binet results. All children who have language handicaps, speech defects, or defective vision or hearing are given the Pintner-Paterson Performance Tests of Intelligence.[1] Data obtained from the so-called intelligence tests are not considered *final* estimates of the child's ability. *They furnish a point of departure for further study of the child.*

c. Educational attainment. All children are given special educational tests when specific inabilities are reported or observed. For this purpose, the following tests are used most frequently:

1. *Stanford Achievement Examination.*
   Test 1—Reading—Paragraph Meaning.
   Test 2—Reading—Sentence Meaning.
   Test 3—Reading—Vocabulary.
   Test 4—Arithmetic Computation.

[1] Additional time (above 3 hours) is allotted to such work.

2. *Buckingham Revision of Ayres Spelling Scale or 7 S Spelling Scale* or special tests taken from text book used (i.e., Horn-Ashbaugh lists).
3. *Gates Reading Tests.*
4. *Compass Diagnostic Tests in Arithmetic.*

d. Nervous stability. Tendencies toward nervous instability are revealed by careful observation of the child during the examination. As a prerequisite to admittance to the clinic training, a special course, The Non-Typical Child, is required of each student. In the course especial attention is devoted to the unstable child whose difficulties are occasioned largely by *malfunctioning* of mentality rather than by variation in the quantitative amount of ability. Some of the most pertinent literature in reference to diagnosis and remedial treatment of unstable children is studied. The students who enroll in the clinic course are familiar already with the major symptoms and treatment of mental instability; the literature is again reviewed and evaluated.

During the various examinations, the student identifies and records any conspicuous tendencies toward instability which are displayed by the child. These tendencies are revealed both through the symptomatic responses to the various examinations and through the general behavior of the child under observation. The following are examples of such symptoms: range upon Binet, bizarreness of response, tendency toward preservation, flightiness, automatic or sound reactions, general nervousness, excessive giggling, etc.

The data in regard to the physical, mental, educational, and emotional development of the child are recorded upon a clinic form report. These data afford the basis for an interpretative account of the status and needs of the child. This account is designed to be of maximum aid to the teacher and supervisor. The end of this work is to secure a more satisfactory adjustment of the child to the school. Three copies are made of the form report as well as the interpretative account. One copy is sent to the superintendent, one to the school supervisor, and one is filed in the clinic.

During the years 1918–1928, 585 boys and 464 girls, or a total of 1049 subjects,[2] have been referred to the clinic for examination and diagnosis. This number does not include two special surveys: one of 458 pre-school children, and another of 100 superior children of IQ 140 or above.

Only since 1924 has the clinic made use of the Porteus Tests, the Pintner-Paterson Scales, the educational tests, and the method of diagnosis of psychopathy outlined above. The record of these phases of the study does not include therefore the total number of subjects reported to the clinic during the ten years period of its existence.

The results of the physical examination are not presented, since differences and standards used in the clinic at various times invalidate the findings. This paper presents: (*a*) the results of tests of verbal and motor intelligence; (*b*) like results of the educational tests; (*c*) a summary of the results of diagnoses of emotional stability; and (*d*) the results of the tests given in the pre-school center.

VERBAL AND MOTOR INTELLIGENCE

1. The 1049 clinic cases referred to the clinic from 1918 to 1928, include 585 boys and 464 girls. The median IQ of the 585 boys is 97.1, S.D. 22.8 and of the 464 girls, 98.1; S.D. 24. The median IQ of 1049 clinic cases is 97.8; S.D. 22.5.

The subjects ranged in chronological age from 4 to 17 years, and in IQ from 32 to 178. Although the intelligence levels of the sexes were approximately equal (averages), there was a slightly greater variability on the part of the girls as compared with the boys who were sent to the clinic.

2. The clinic cases are on the average only slightly inferior to the group of 1000 unselected children studied by Terman. When the negroes of the clinic group are excluded, the median IQ of the subjects is 100, S.D. 21. For Terman's group of unselected children the median IQ is 100.5, S.D. 13.

[2] Mrs. Mary Chord Hazard assembled data regarding about 700 of the subjects. These subjects were ones observed by the student examiners. The remaining cases were examined by the writer.

3. The clinic cases include 7.5 per cent more potentially feeble minded, 5.4 per cent more very superior, and 3 per cent more subjects of "potential genius" rank than the unselected group studied by Terman. The unselected group has 24 per cent more "normal" subjects than the clinic group. A conspicuous difference between the two groups lies in the greater frequency of extreme cases in the clinic group.

4. The median mental age of boys on the Porteus Test is 10.8 years with S.D of 2.14.[3] The median chronological age of the same group of boys is 10.5 years with S.D. of 2.08; the median of the girls on the Porteus Test is 10.2 years with S.D. of 2.87. The median chronological age of the same group of girls is 9.9 years with S.D. of 2.7. Both boys and girls rank slightly above the averages upon the test.

5. The coefficients of correlations between Porteus mental age and Stanford-Binet mental age for the boys and the girls are respectively $+.642$ P.E. $\pm.04$ and $+.78$ P.E. $\pm.018$. The Pearson Product Moment Formula was employed in calculating these coefficients.

THE EDUCATIONAL TESTS

The clinic cases are from all the school grades with the fifth as a median, the range being from kindergarten to university. With the aid of Strayer's Standard of Gradation, a study was made of the chronological age grade progress and mental age grade progress of 751 subjects from grades I to VIII inclusive. The results of the various educational tests yielded additional data.

1. Approximately one-third (33.5 per cent) are of normal chronological age for their grade placement, 30.8 per cent are above average, and 35.7 per cent are underage. The children as a group seem to be placed in the various grades according to their birthday ages.

2. Only slightly more than one-fourth (26.3 per cent) are of

[3] All children were not given the Porteus test since the test appeared too easy for those of advanced C.A. Coefficients of correlations were calculated for 200 boys and 200 girls only.

normal mental age for their grade placement, 28 per cent are above average, and 45.7 per cent are underage. These results are significant; the children sent to the clinic in recent years appear to possess restricted mental ability; ability which makes normal functioning in their respective grades difficult.

3. The results obtained by giving the educational tests have varied from year to year;[4] at the present time, there is a tendency to use diagnostic tests rather than the general educational surveys.

Eighteen per cent were rated at grade according to the standardized educational tests, 57 per cent were rated accelerated in grade placement in advance of their actual educational attainment, and 26.3 per cent were rated retarded in grade placement as compared with educational attainment. More than half of the children, therefore, appear to be functioning below the standards of the grades in which they are placed.

Further study of the educational test data is illuminating. Reading is the subject in which the poorest educational attainment was elicited according to the results of the educational tests. Only 20 per cent of 250 girls reached or exceeded the median for their grade in this subject and only 14 per cent of 250 boys reached or exceeded the median. These results were secured by comparison of the childrens' test results with the norms given in the Stanford Achievement Examination. There were 22 per cent of the girls and 28 per cent of the boys who could be characterized as problem or disability cases. A case study which follows later shows the type of remedial endeavor which is instigated for those seriously handicapped in reading

The subjects were found to be functioning far below standards when the Buckingham Ayres Scale of Spelling Ability was administered. Their achievement, however, upon the Stanford Achievement Dictation Exercise (Spelling) Test 9 excelled the Buckingham Ayres' rating. Spelling attainment (as measured by this scale) averaged one year below grade placement for the

[4] Different standardized tests have been used from time to time.

group. The children approached closely the goals set for them in the Horn-Ashbaugh tests.

The children were decidedly restricted in the types of arithmetic ability measured by the Stanford Achievement Tests (4 and 5). The average for the entire group was one year below grade placement upon Arithmetic Computation and one and one-half years below grade standing upon Arithmetic Reasoning.

EMOTIONAL STABILITY[5]

The interpretative reports of the graduate students are the source for the data here presented. These reports record special or unusual responses in the behavior of the subjects. They are too individualistic and subjective in character to be summarized in all details, but an effort has been made to interpret the reports of nervous stability. This type of report has been used only for the subjects referred to the clinic since 1924.

1. Since, as explained in the introduction, the examiners are somewhat capable in diagnosing nervous instability, their judgment is relied upon in determining psychopathic tendencies, and no further delineation is made herewith of the Stanford-Binet Test to discover bizarreness of response, range, etc.

"Slight tendency" includes notations such as: "predisposition," "some evidences," "a few indications," and the like. "Marked tendency" includes: "very unstable," "marked instability," and the like. The largest per cent of psychopathic tendency is in the group of IQ 32–69, where 47.3 per cent were thus affected. No such cases are recorded in the group of highest intelligence, 140–178 IQ, with the next smallest per cent in the group of IQ 110–139 (4.8 per cent).

Sixteen and six-tenths per cent showed tendencies to instability. The National Committee upon Mental Hygiene estimates that 3.5 per cent of the school population is psychopathic. According to this estimate there were almost five times as many

[5] This part of the study deals with the subjects studied by Mrs. Hazard.

psychopaths in the clinic group as in an unselected group. The writer is cognizant of the chance of error in the data interpreted by the students.

RESULTS OF THE PRE-SCHOOL CLINIC

The special study of pre-school children was made by regularly qualified graduate students and the writer during the months of April and May of the years 1926 and 1927 and by the writer during the years 1928–1929. Usual development histories and physical test records were secured. Each child between the ages of three and five was examined mentally by means of the Stanford Revision of the Binet-Simon Intelligence Test. There were in all 458 complete records obtained from 257 boys and 201 girls. The range in chronological age was from 3 years to 6 years; the mental ages ranged from 3 years to 8 years; the intelligence quotients from 56 to 153.

The following phases of the study are here summarized: (1) The relationship between the mental ability of the children and occupational status of the parent; (2) the relationship between sex and mental ability; (3) the correspondence between mental ability and certain indices of growth; (4) an analysis of the difficulty of certain parts of the Stanford-Binet Test.

1. The occupational status of the parents of the exceptionally capable and exceptionally retarded children was determined by reference to the classification employed in the United States Government Census Reports. There were ten times as many children of professional parents as of industrial parents classified in the group of IQ 135 and above; there were five times as many children from the industrial class as from the professional class in the groups of IQ 85 or below.

The degree of relationship between mental ability and occupational level was determined by correlating the Binet mental ages with the probable error values for various occupations listed in the Barr scale. The Barr scale assigns values to various occupations according to their estimated degrees of difficulty (evidenced through the pooling of judgments). One hundred occupations, each definitely and concretely described, were

rated by 30 judges according to the grade of intelligence which each was believed to demand. The ratings were distributed and probable error values assigned.[8]

The coefficient of correlation between M.A. and P.E. value on the Barr scale is +.48 P.E. ±.02. It is clear therefore that a substantial positive relationship exists between occupational level of the parent and mental ability of the child. It must be remembered that these data do not preclude the possibility of mentally superior children being found in homes of industrial workers; superior children are to be found therein but their frequency is less than in the homes of professional parents; they are however much less frequently encountered than in the homes of industrial workers.

2. In the Stanford-Binet Test the girls made both the highest and the lowest scores, thus having a greater range of scores than the boys. The average performance of the girls (Median IQ 116) exceeds that of the boys (Median IQ 108) by 8 points. The superiority of the girls may be accounted for at least in part by the fact that among the girls there was a larger percentage with fathers in the professional and a smaller percentage in the industrial groups. No conspicuous sex differences were found when a comparison was made between the median intelligence scores assembled according to the occupation of the parents.

3. The height and weight of each child were recorded and compared with the Baldwin-Wood Norms to determine the ratio of each of these measurements to that of a normal child of the same age and sex. The height and weight quotients thus established were correlated with the intelligence quotients with the following results:

	Boys	*Girls*
Correlation between IQ and HQ	+.17 ±.04	+.16 ±.04
Correlation between IQ and WQ	+.15 ±.04	+.16 ±.04
Correlation between HQ and WQ	+.71 ±.02	+.64 ±.02

[8] "Genetic Studies of Genius," Vol. I. Terman et al. p. 66 to 68, Stanford University Press, 1925. It was necessary to employ considerable subjective judgment in assigning values to various occupations which did not fall exactly in the 100 listed. Three judgments were utilized in this endeavor.

The relationship between physical development and mental ability was determined also by correlating mental age with height and weight respectively with chronological age held constant. The coefficient of correlation between height and mental age +.15, between weight and mental age, +.14. Collectively these data disclose a positive but almost negligible relationship between mental and physical maturity (of the types measured).

4. The difficulty of the various items of the Stanford-Binet Test was ascertained by analysis of the responses of the pre-school children. The successes and failures of representative children upon the various items of the test were assembled and considered. In year III test 6 was found to be the most difficult; 19 per cent of the children missed this test.

At year IV tests 4 and 6 proved decidedly difficult; only 35 per cent of the four year old Lawrence children passed test 4, and thirty-two per cent test 6. At year V, tests 5 and 6 proved much less difficult for both four and five year olds than tests 4 and 6 in year IV.

These are the tests most obviously in need of regrading. On the whole, the Stanford-Binet Examination does not serve adequately in testing pre-school children.

Table 1 presents in detail an analysis of the test results.

The following examples of interpretative reports indicate the nature of the work of the clinic.

Case I

Name: J. Grade 4A. C.A. 10 years (126 months). M.A. 132 months. IQ 105.

J. is a rather superior child whose ability probably is not adequately represented in the IQ obtained from the Stanford-Binet Test, since he so nearly passed several tests which must be graded failures. He is quiet, serious, and very coöperative. The physical tests show J. to be normal in every regard.

On the Stanford-Binet Test J. displayed a ten year vocabulary, which is normal for his age. He failed, however, in the ten year reading test. He made more errors than are allowed

TABLE 1

Showing the successes of 458 children of the Lawrence Pre-School Clinic in the several tests of the Stanford-Binet Test expressed in per cent

YEAR AND TEST		3-YEAR-OLDS KANSAS	4-YEAR-OLDS KANSAS E	5-YEAR-OLDS KANSAS S
III	1	100	97	99
	2	100	99	99
	3	91	98	99
	4	86	98	99
	5	91	95	97
	6	81	97	99
IV	1	62	86	86
	2	33	84	93
	3	24	50	81
	4	10	35	81
	5	48	83	93
	6	14	32	65
V	1	38	72	94
	2	15	39	71
	3	28	35	83
	4	14	50	85
	5	5	62	89
	6	28	38	91
VI	1	5	37	57
	2	5	25	49
	3	5	12	32
	4	5	31	64
	5	0	25	46
	6	5	17	42
VII	1	5	0	15
	2	1	2	11
	3	5	5	14
	4	0	1	8
	5		0	11
	6		0	7
VIII	1		0	1
	2		0	0
	3		2	1
	4		0	6
	5		3	7
	6			0

for success in oral reading and his rate of reading was decidedly slow. It was obvious that many of the failures in the entire Binet Test were due to faulty reading habits.

The feeling that restricted reading ability prevented an adequate display of mental ability was corroborated by J's scores upon performance tests. On the Porteus Test J. secured a rating of 13 years, his maximum performance upon any test. His median mental age upon the Pintner-Paterson Test was 12 years, 6 months. One may conclude therefore that J. is mentally a somewhat superior child, one who is seriously handicapped by faulty reading habits.

The Stanford Achievement Test yielded the following ages: Paragraph meaning, 8 years, 2 months; Sentence Meaning, 9 years, 1 month; Word meaning, 9 years, 4 months.

J. succeeded moderately well when he dealt with vocabulary tests or with sentence meaning tests. When the test unit was larger than the sentence, he was unable to cope successfully with the situation. Particularly noticeable was the fact that J. is a word reader. He is capable of interpreting paragraphs only as does an average second grade child. It is desirable therefore that attention be directed toward developing habits of attacking the paragraph as a whole. The analytic-synthetic method of teaching reading might well be used in correcting the weakness. Individual remedial work is of course essential. The technique suggested by Buswell and Wheeler might be utilized in remedial endeavor. Coördinate development of vocabulary is essential also if best results are to be secured. (A graduate student is now devoting 30 minutes daily to J. in an attempt to establish correct reading habits. The remedial work is proving unusually successful.)

Case II

R. has been under clinical observation for several years. R. was brought to the clinic in March, 1926. His teachers reported that he could not learn to read and "although he seemed bright enough might be feeble-minded." His mother is an intelligent, coöperative person. She reported that R. "was bright when compared with other children, but simply could not read."

The physical tests showed R. to be normal in several regards; he was apparently in excellent health. Neither visual nor auditory defects were revealed. The developmental history revealed nothing unusual except inability to read and slight nervousness.

The Binet Examination of intelligence yielded a mental age of 86 months; R's chronological age was 91 months; the resultant intelligence quotient, 95. Obviously R. has sufficient general ability to succeed fairly well with the reading function.

Several features of the examination were noticeable. R's range on the Binet Test was 5 years. He exhibited tendencies toward perseveration, he gave sound reactions to words, and displayed extreme excitability. His attention was difficult to secure and hold. R. was very loquacious, a tendency to talk excessively and irrelevantly was a constant feature of the examination.

R. was in grade 2A; therefore second grade attainment scales were first employed in examining him. The Stanford Achievement Examination yielded a zero score, the Gray Oral Reading Test, zero, Buckingham Ayres Spelling Scale, zero (all lists). The Pressey Attainment Scales for first grade children were given. On these also R. was unsuccessful in making positive scores. An interesting defense was exhibited by the child. In the Spelling test, he complied with directions, writing something always in response to the prompting but never once spelling a word correctly. On the reading test, he underlined words in response to directions, always working by chance. It was found that two letters only could be recognized in isolation and no word could be identified.

Individual remedial work was deemed essential. Since demotion appeared inadvisable a University student was selected to give R. 30 minutes of individual work daily. Remedial work began March 30 and continued until June 15.

The following factors were considered important ones in effecting R.'s failure in school work, especially in reading:

a. Nervous instability, flights of attention and lack of application.
b. Lack of adequate number of sight words for reading.

c. Lack of recognition of many of written or printed letters and all of printed capital letters.
d. Lack of knowledge of sounds of letters, phonic units, or means of attack on new words.
e. Habits of guessing instead of thinking.
f. Inability to function at 2A level with the consequent development of numerous defense mechanisms.

Since the method of teaching reading which employs the word or larger unit of perception had proved unsuccessful with R., the phonogram and the letter were taught as units of perception. Accordingly, the list of phonic elements which are usually taught in the first grade was presented. These were selected from the list compiled by Washburne, Jaycox, and Vogel.

These units were presented in the following manner. Words from word families were pronounced to R.; he was asked to observe the common element. Practice ensued in pronouncing families of words with subsequent identification of the common element. Recognition of the phonic units was then made automatic through the use of flash cards to which R. responded by pronouncing the unit displayed. Practice was given repeatedly in pronouncing words in which the known phonogram occurred.

A basic list of sight words was taught through flash-card exercises (68 in number which the school had used, plus 26 which were made to meet R.'s individual needs). These were displayed and the customary responses elicited.

Phonetic words were taught by having R. sound them out; new words were first divided into known elements. R. was instructed to identify and pronounce the common element. Unphonetic words were taught as sight words; in these, the letters were attended to and in many cases sounded. Spelling was taught by leading R. to identify and pronounce the phonic elements occurring in new words.

After the sixth week of remedial work, an attempt was made to deal with the phrase or sentence as a unit by using the Buswell-Wheeler phrase cards which require dramatic action or pretense. Very little was accomplished however at this time.

After the eight weeks of remedial work, the Gray Oral Reading Test was given, as well as Lists E, F, and G of the Buckingham Ayres' Spelling Scale with the following results:

Gray Oral Reading Test Results

Paragraph	I	II	III
Errors	4	2	5
Time for reading	57 sec.	70 sec.	107 sec.

Buckingham Ayres Spelling Scale Results

	R's score	*Standard for 1st year*
List E	76	92
List F	85	88
List G	66	84

At this time R. recognized most of the words taught immediately upon sight. Some verbs still caused confusion. His knowledge of phonics seemed to enable him now to attack new words without guessing.

The remedial work with R. was continued through 8 weeks of the summer under the direction of another graduate student of the University of Kansas. The earlier phonetic drill was carried forward, the list and technique suggested by Washburne, Vogel, and Jaycox were utilized. Spelling was now taught according to the principles set forth by Horn. The Buswell-Wheeler Readers and technique were employed in reading. Noticeable improvement was elicited.

In September, 1926, R. again entered the 2B grade. During the academic year, the remedial work was continued by R.'s classroom teacher. In March, 1927, one year after the remedial measures were instigated, the following test results were secured:

1. Gray Oral Reading—R's score equal to high 2nd grade scores.
2. Pressey First Grade Reading Scale.
   (Attainment Scale, Form A)

	Scores	*Tentative Norms grade 1A*
Word Test	24	18.5
Sentence Test	15	13.0
Total Score	39	31.0

3. Buckingham Ayres' Scale.

	R. Per cent	Norms Per cent
List F	92	88
List G	78	84

4. Gates Primary Reading Test.

Type 1. Word Recognition (Grades 1 and 2, Form I). R. made a score of 35. This gives a reading grade 2.7 and a reading age 8.2 years.

Type 2. Words, Phrase, and Sentence Reading. R. made a score of 12. This gives a reading grade 1.90, and a reading age 7.4.

The above tests show that R. has improved remarkably during the past year. He still has many difficulties to overcome if he is to keep up with his class. He has passed the first grade standards in the Pressey Achievement Test and the Word Recognition and the Gates Primary Test. R. is in the second grade and is with children somewhat farther advanced than he in the reading function.

It is obvious that if any lasting good is to be secured from the remedial measures employed thus far, R. must have individual remedial attention directed toward continued increase of vocabulary and development of the ability to comprehend large units. (This work is now being carried on by a graduate student in the University of Kansas.[7])

Case III

Name: V. Grade 6A. C.A. 176 months. M.A. 113 months. IQ 64.

V., 14 years, 8 months of age, has a mental age of 9 years, 5 months, and a resultant intelligence quotient of 64, which classified him as potentially feeble-minded. His amiable disposition and willingness to attempt a task make his mental deficiency less noticeable at first. His responses to the Stanford-Binet Test were uniformly slow, labored, and inadequate.

[7] In March, 1929, R. was sent again to the clinic. He is now in grade 4A and is happy and successful in his school work. He now is able to reach 4A standards upon such tests as the Gates, Stanford Achievement, etc.

He showed no desire to criticize his achievement; he accompanied most of his replies with a perfunctory question as to their correctness. His lack of ability to associate ideas was particularly marked on the free association test, in which he named but twenty-nine words in three minutes (standard for 10 years is 60 in three minutes). Only visible objects were named.

V.'s age on the Porteus Maze Test of motor coördination was the same as that of the Binet Test, 9 years. In reading, V.'s achievement was that of the fourth grade child in rate, but he reached only third grade quality in comprehension. His composite score on the Stanford Achievement Test, Form A, was that which one would expect of a 4th grade child at the beginning of the academic year. There was a general inferiority displayed on all parts of the examination.

This child is unable to do average school work. His inability to comprehend reading material makes his present classification (sixth grade) difficult for him and for his teacher. He might well be placed in a special class and given work suited to his mental ability. As stated at the first of his report, V.'s disposition to try, and his tendency to be happy and satisfied, probably make his handicaps less noticeable, but his limited capacity and educational attainment do not justify his present grade classification.

The case studies given above indicate the nature of the diagnoses made in the clinic. The question is often asked: "How do the teachers, parents, and public school officials look upon the work of the Clinic?" The school and community have responded wholeheartedly to the work. It is impossible for the clinic now to provide observation for the numerous cases seeking help. The following excerpt from a letter written by Miss Susan B. Friend, Grade Supervisor, is indicative of the reception accorded the efforts of the clinic:

"This report would not be complete if some reference were not made to the effect the work of the clinic has had both on parents and on the teachers. Parents who formerly were antagonistic to tests and testers now welcome any suggestions from the clinic as being veryworth-while and teachers have come to recognize the significance of individual instruction and to realize the importance of the emotional factor in the child's life."

AN ENQUIRY INTO THE STANDARDIZATION OF THE KOHS' BLOCK-DESIGN TEST[1]

A. JEAN BROWN, M.A.

Psychologist, Mental Hygiene Division, Department of Public Health, Toronto

There is an urgent need for single performance tests or sets of tests for clinical use. A careful survey of available performance tests had indicated that few show consistent age or grade increments over any period of years. The wide adoption of the age level concept makes it advisable to select a test that conforms to this criterion.

This article reports a study of the Kohs' Block-Design Test carried out by the writer in the Toronto Public Schools in 1927–1928. This test shows an increase of the median score at successive age levels. The equipment is not complex, the instructions are brief and the test requires a moderate amount of time to give. Successful performance, speed and accuracy are combined in the total score. The weight ratio of these factors is roughly 4:2:1. The method used was that described by Kohs (2). It was found necessary to give the trial design to the subject before beginning the test. The coloured designs and blocks rarely failed to motivate the child to action—seldom was urging required.

In the selection of cases three steps were taken. First, the writer examined the files of the Division and located two schools in which the distribution of intelligence test scores was approximately that of the average for the City. Secondly, three age levels were taken so that a large number of cases would be

[1] This study is based on tests carried out by the writer in the capacity of psychologist on the staff of the Mental Hygiene Division. Director of Mental Hygiene Division, E.P. Lewis, B.A., M.B. Medical Officer of Health, C. J. O. Hastings, M.D., L.L.D.

tested at each age level. The ages chosen were, nine, ten and eleven because below nine years some children are not in school and above eleven few of the brightest children have been transferred from Public Schools to Collegiates. Thirdly, all children who were within one month of birthday were examined, irrespective of grades in which they were found. It is believed that the cases tested may be taken as a true sample of the school

TABLE 1

Showing age and sex distribution of cases examined

AGE	GIRLS	BOYS	TOTAL
9	21	21	42
10	32	19	51
11	27	34	61
Total.....	80	74	154

TABLE 2

Showing distribution of time scores for each age level

AGE	SCORE											
	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–99	100–109	Total
9	16	10	6	5	2	2			1			42
10	15	10	9	5	4	5	3					51
11	7	15	13	9	4	3	4		2	3	1	61
Total	38	35	28	19	10	10	7		3	3	1	154

population of Toronto of these ages. Table 1 gives the distribution of cases at each age level and for each sex.

DATA COLLECTED

The time required by each pupil to complete each design was noted to the nearest second and recorded on a specially prepared blank. A group intelligence test was given to each pupil. (National A. Form I, for grades 3 to 8. Pintner-Cunningham Primary for grades 1 and 2.)

RESULTS

Time scores were calculated on an age and grade basis. Table 2 shows the distribution of time scores for each age level.

There is a wide range of scores at each age level and marked overlapping. Both the median and the average scores with the average deviations were calculated for each age level. See table 3.

The average scores which were obtained in this investigation more closely approximate Kohs' smoothed median scores than do our median scores.

The correlation of the Block-Design mental age and the National Intelligence mental age is +.46 ±.04 (104 cases).

TABLE 3

Showing median and average scores at each age level

	AGE		
	9	10	11
Kohs' Norms	18	25.5	36
Median Score	16±13	20±15	27±18
A.M. Score	19±13	24±15	34±16

This correlation is positive but it is not high. It may be noted that the National Intelligence Test in addition to being a language test involves reading to a great extent. Dr. Kohs reports that the Block-Design scores correlate lower with reading and language than any other school subject.

CONCLUSION

The point which seems of most importance in this investigation is the wide range and consequent marked overlapping at each age level. It is true that the median scores increase at each age level and our arithmetical means closely approximate the medians given by Kohs, but when these averages are interpreted in the light of their arithmetical deviations the validity of the test is called into question. It is recognized that Kohs found a correlation of +.84 ±.01 with Binet mental age.

Judged by the principle of demarcation of age groups, however, the validity of the Kohs' test is of doubtful clinical value.

BIBLIOGRAPHY

(1) Bronner, A. F., Healy, Wm., Lowe, G. M., Shimberg, M. E.: A Manual of Individual Mental Tests and Testing. Judge Baker Foundation, Publication No. 4.

(2) Kohs, S. C.: Intelligence Measurement, A Psychological and Statistical Study based on the Block-Design Tests. Macmillan Company, 1923.

(3) Kohs, S. C.: The Block-Design Tests. Jour. Exper. Psychol., 1920, Vol. 3, 357–376.

AN EMPIRICAL STUDY OF THE EFFECTS OF GROUPING DATA IN CALCULATION OF R BY THE PEARSON PRODUCTS MOMENT METHOD

ALVHH R. LAUER
Ohio State University

INTRODUCTION

The increased use of statistical methods and the development of machines for rapid calculation of results have caused many short cuts in computation to be introduced or adapted from older methods. It is the purpose of this paper to present data on certain aspects of such issues. Complaints have been raised against statistical results generally by well known writers. No less an authority than Spearman (7) has voiced a note of warning, "The experience of the writer has convinced him that there is only too much truth in the recent statement by a prominent psychologist, that few of the published correlations are safe from downright errors of calculation. Time and again, grave mistakes have been detected, and this even in correlations published by authors who stand in high repute both as mathematicians and as careful workers."

It is the writer's contention that the indiscriminate application of formulae is a source of error which may account for certain mistakes attributed to faulty computation. Most authorities realize limitations in the application of specific principles and formulae, as we note from Thurstone, (9) "Of course, it is not to be expected that any one method of scaling will apply to all types of educational and psychological test material, and I am not attempting to offer a single solution for all scaling problems."

Orleans (4) has developed a method for calculating a coeffi-

cient of correlation without plotting by a derived formula of the conventional

$$r_{xy} = \frac{\Sigma\, xy}{N\, \sigma_x \cdot \sigma_y}.$$

The formula is written

$$r_{xy} = \frac{\sigma_x}{2\, \sigma_y} + \frac{\sigma_y}{2\, \sigma_x} - \frac{\sigma\, d^2 + Md^2 - (M_x - M_y)^2}{2\, \sigma_x\, \sigma_y}.$$

However he states that the method is sensitive and when not properly used will result in an error in r of from .10 to .20, especially when the means of the two sets differ greatly in magnitude. Research workers often select a formula and apply it to any data which may be under immediate consideration. It is evident that some rather hap-hazard results are likely to occur.

There are several sources of error which have been suggested by writers relative to the practice of grouping data. Kelley (3) warns that it is best not to use iess than twelve groups as a lower limit. Instances have come to the knowledge of the present writer where as few as five groups have been used involving rather large class intervals.

The class interval itself is another item to be considered. Fisher (1) has held that the interval should not exceed one quarter of the standard deviation. His primary consideration at that point, however, was measures of dispersion and not correlation. In this connection the failure to use either Sheppard's correction for midpoint values or Pearson's correction for mean values presents another problem to be analyzed. Also the normality of the distribution is assumed in applying most formulae but is rarely experienced in actual situations.

STATEMENT OF THE PROBLEM

Several such factors were studied by the writer while working over college entrance test materials involving 18 variables, with 302 possible intercorrelations, using 492 cases, at Iowa State College. In an effort to locate the cause of certain dis-

crepancies in r's calculated, a careful analysis was made of the most disturbing cases of correlation encountered.

Since Hollerith machines were available the raw data were punched on cards and the various coefficients computed from these. (Raw scores were used for administrative reasons.) Because of wide ranges in certain test scores, and large figures involved, it was suggested by an assisting statistician that material be grouped according to the nature of the data as needed.

It might be stated that this irregularity in procedure resulted in the data being semi-grouped and semi-coded. For example the values of X were grouped in a consistent manner and the mid-point used as a coded value for the group. The actual Y values falling in each group were summed and multiplied by the mid-point values of the X groups. To get the sum of the products XY the results of the above calculations were totaled. The sum of the X^2s were obtained by squaring each mid-point and multiplying it by the frequency or number of X values in this group. These totaled gave the sum of X^2s. The mean number of groups used was 9, mean class interval of 6 with a mean range of 53.8. The Pearson Product-Moment Method was used, the formula being a derivation of the well known Ayres[1] formula used by Snedecor and Wallace (5).

$$r_{xy} = \frac{\Sigma XY - (\Sigma X)\, M_y}{\sqrt{\Sigma X^2 - (\Sigma X) \cdot M_x} \cdot \sqrt{\Sigma Y^2 - (\Sigma Y)\, M_y}}$$

It is evident that the use of exact scores as individual values makes this a very reliable and valuable formula but it was found that grouping caused some difficulty and that results did not check when the variables were changed around. Further it can be shown that when the X values grouped are all the same this method will check out accurately but in certain instances when otherwise grouped wide discrepancies are likely to result. That this might logically occur is but a matter of simple algebra

[1] Written in different forms by Harris in 1910, by Thurstone in 1917, and again by Ayres in 1921.

since there is no proof that $\left(\frac{a+b+c}{n}\right)(x+y+z)$ is equal to $(ax+by+cz)$. (Here a and x, b and y, c and z are paired values of the X variable and the Y variable respectively and n is the number of cases falling in each group.)

Upon analysis it was noted that grouping scores on X, as described above, did not give the same results as the same data grouped on Y. The coefficient might vary considerably. Further it was observed that the mean of the two r's obtained was quite close (within the probable error) to that obtained by using individual scores without grouping. Consequently all r's used were calculated from both sides of the table and out of the possible 302 combinations 41 were thoroughly checked for accuracy and used in calculating the results given below. This represented actually 82 correlations as each was done twice.

TABLE 1

Difference in r's	Range	.0011 to .2457
Difference in r's	Mean	.1025
Difference in r's falling within	1 P.E. of r	14.2%
Difference in r's falling within	2 P.E. of r	7.3%
Difference in r's falling within	3 P.E. of r	24.9%
Difference in r's falling within	4 P.E. of r	6.9%
Difference in r's falling outside	5 P.E. of r	47.0%

RESULTS

In consideration of these discrepancies it was found that they were entirely too large to be accounted for by chance. The significance of the differences between the two r's so obtained on the same data will be indicated in table 1.

Various factors were studied in further analysis of data. Orleans (4) states that his formula is valid only when the means of the two sets are approximately equal. Kelley (3) recommends more groups and Fisher (1) thinks the interval must be relatively small. Since the normality of the distribution is assumed in using the variations on the Pearson products moment formula, skewness was also taken into consideration.

A high correlation would be expected between certain of the factors studied due to spurious correlation resulting from inher-

TABLE 2

Zero order coefficients between discrepancy and other variables

$r_{D C_D}$	$= +.5112 \pm .08$	$r_{D M_d}$	$= +.2183 \pm .10$
$r_{D SD_d}$	$= +.4817 \pm .08$	$r_{D r}$	$= +.0400 \pm .11$
$r_{D S_s}$	$= +.4310 \pm .09$	$r_{D \frac{I}{SD}}$	$= +.0385 \pm .11$
$r_{D C_m}$	$= +.3100 \pm .09$	$r_{D G_m}$	$= -.0229 \pm .11$
$r_{D S_d}$	$= +.2729 \pm .10$	$r_{D R_d}$	$= +.0087 \pm .11$
$r_{D I_m}$	$= +.2773 \pm .10$		

SD_d = Difference in the standard deviations.
D = Discrepancies in r's.
C_d = Difference in coefficients of variability for each of the two sets.
C_m = Mean of coefficients of variability for the two sets.
S_s = Sum of skewness as measured by the formula $\frac{3\ (M - Mdn)}{S\ D}$. Skewness of one set added to skewness of the other.
S_d = Difference of skewness as measured by the above formula—one subtracted from the other.
r = Coefficient of correlation.
R_d = Difference in range of two sets.
I_m = Mean class interval used in the sets.
M_d = Difference in mean of two sets.
$\frac{I}{SD}$ = Ratio of class interval to the standard deviation of the same variable (suggested by Fisher).
G_m = Mean number of groups used in each of the two sets.

TABLE 3

Intercorrelations between five significant factors

	I_m	C_d	M_d	SD_d
S_s	$+.4166 \pm .09$	$+.2566 \pm .10$	$+.7185 \pm .05$	$+.4922 \pm .08$
SD_d	$+.6912 \pm .05$	$+.2637 \pm .10$	$+.3811 \pm .09$	
M_d	$+.3765 \pm .09$	$+.1625 \pm .11$		
C_d	$+.5186 \pm .08$			

ent relationship. However, it was only desired to trace down as far as possible the relative causes of discrepancies. For this

reason the absolute values of the results are of secondary importance. Correlations were computed between the data in table 2 by reversing variables, i.e., by working them from both sides of the table. Only the products moment method was used.

A multiple correlation was now run between D as a criterion and the five variables given in table 3. This gave R = +.7733 when corrected for number of variables and cases by the formula (see 5).

$$R^2 = 1 - \frac{1 - R^2}{1 - \frac{M}{N}}$$

gives R = +.7330, when M = number of variables and N = the number of observations. The regression equation is as follows:

$$\overline{D} = .2526\ C_d + .0129\ SD_d + .0822\ S_s + .0255\ I_m - .0005\ M_d.$$

An interpretation of the relative contributing influences of these three variables will depend upon which device is used as a criterion. If β values are used, the order of descending merit is; (a) differences in the relative variability of the two sets, (b) amount of skewness, and (c) the differences in absolute magnitudes of the standard deviations. The regression equation will give somewhat different results as indicated above. It suffices to say, from these data, that certain statistical errors are likely to occur when grouping or coding schemes are applied to sets differing with respect to the above three items.

CONCLUSIONS

The results seem to indicate:

1. That a semi-coding scheme is not good and produces at times certain inconsistent results.

2. That these inconsistencies or discrepancies are due to several factors, the most pronounced being:

a. Differences in variability of the two sets.

b. Amount of skewness or departure from normality in the distribution of one or both of the sets.

c. Difference in standard deviations of the two sets of scores.

3. Other factors which have little influence which may operate in and of themselves:

a. Size of class interval.

b. Number of groups used.

c. Ratio of standard deviations to the means.

d. Size of r's calculated, i.e., high or low correlation between variables concerned.

e. Difference in the ranges of the two sets.

4. Still other factors which have not been analyzed in this study but were suggested by it as important items to be considered in any grouping or coding scheme:

a. Arrangement of data into groups such that certain periodic groupings of scores in bimodal or multimodal distributions are properly equalized; e.g. in handling grades there appears a general tendency for teachers to give more 75, 80, 85, etc., than 76, 81, or 86, etc. These must be properly distributed so as to equalize the effect on correlation.

b. The treatment of isolated cases as individual rather than group values when such would artificially increase the range.

c. Extreme care in deciding class interval limits.

d. That certain data grouped in the same or different groups by a slight variation in method may give different results. Neither Sheppard's correction for midpoint values nor Pearson's correction for mean group values seem adequate or generally applicable.

5. In general grouping is likely to cause difficulty and should be resorted to only when the range and number of cases make the saving of time more important than extreme accuracy of results. It is assumed that these findings hold only when the method of grouping described above is used but it is suggestive of other possible sources of error. Further investigation should

be made of the effects of grouping and coding as is customarily done with published correlation charts.

REFERENCES

(1) Fisher, R. A.: Statistical Methods for Research Workers. Oliver & Boyd, London, 1925, p. 52 et segue.

(2) Garrett, H. E.: A Modification of Tolley and Ezekiel's Method of Handling Multiple Correlation Problems. Jour. Educ. Psychol., 1928, xix, pp. 45–49.

(3) Kelley, T. L.: Statistical Method. Macmillan Co., 1923, pp. 50 and 167–171.

(4) Orleans, J.: Correlations without Plotting. Jour. Educ. Psychol., 1927, xviii, pp. 310–315.

(5) Snedecor and Wallace: Correlation and Machine Calculation. Iowa State College Bulletin, 1925, pp. 8 and 47.

(6) Rugg, H. O.: Statistical Methods Applied to Education, Houghton Mifflin Co., 1917, pp. 147–180.

(7) Spearman, C.: The Abilities of Man. Macmillan & Co., 1927, p. 157.

(8) Symonds, P.: Variations of the Product-Moment (Pearson) Coefficient of Correlation. Jour. Educ. Psychol., 1926, xvii, pp. 458–469.

(9) Thurstone, L. L.: A Method of Scaling Psychological and Educational Tests. Jour. Educ. Psychol., 1925, xvi, no. 7, p. 433.

NEWS AND NOTES

A new national educational periodical, the *Junior College Journal* will begin publication in October, 1930. It will be published by Stanford University Press under the joint editorial control of the American Association of Junior Colleges and the School of Education of Stanford University. Dr. Walter C. Eells, of the Stanford University School of Education, will be editor-in-chief, with Doak S. Campbell, of Nashville, Tennessee, secretary of the American Association of Junior Colleges, as associate editor. The new publication will be devoted to the interests of the more than four hundred junior colleges in the United States which heretofore have had no periodical devoted especially to their interests.

Stanford University announces the appointment of Dr. Heinrich Kluever, of the Institute for Juvenile Research, Chicago, as a member of its staff during the coming summer quarter.

Smith College School for Social Work announces a quarterly publication, The Smith College Studies in Social Work, the first number of which will appear about October 1, 1930. The editors, Professor Everett Kimball and Dr. Helen Leland Witmer, plan to publish the results of clinical and social investigations carried on by the students in the fulfillment of requirements for the degree. The subscription price will be two dollars a year.

Yale University announces a new plan of study in the Department of Education to be introduced next fall by which a group of graduate students working for the M.A. or Ph.D. degrees and limited to fifty in number will be selected with especial reference to their intellectual maturity and capacity for independent thinking. This group will be given freedom from conventional courses and credit requirements and will be completely separated from extension or undergraduate students and will have at its disposal practically the full time of seven members of the faculty. A general seminar will hold two-hour sessions each week which will help the student to achieve a unified understanding of the whole educational program, including the interrelation of its parts, and will be based on a syllabus outlining the fundamental data and problems of educational philosophy, history, psychology, sociology, curricula, learn-

ing, instruction, supervision, organization, administration and financial support.

According to its eleventh annual report the Commonwealth Fund made appropriations totalling $2,028,759.26 during the year ending last Septemper 30. Among the year's developments were the opening of two new rural hospitals and the adoption of a new public health program in this country; the establishment of a child guidance clinic in London; and the completion of the Fund's child health program in Austria. Fellowships for graduates of British universities to study in America were increased from twenty-five to thirty-five. The largest special grant made to an outside organization was one of $214,500 to the University of London for the establishment of a chair of American history. In the establishment of a number of rural hospitals some 410,000 people are able to enjoy hospital service which in former years they had to seek at a distance. The objectives of this program besides providing these facilities are to attempt to raise the level of rural health through stimulating the medical profession and the public, to develop practicable public health services suitably linked with the hospitals and to foster an intelligent interest in preventive medicine.

BOOK REVIEWS

H. Meltzer, and E. M. Bailor. *Developed Lessons in Psychology.* Harcourt, Brace & Company, 1929, pp. x + 199.

H. H. Remmers, and G. C. Brandenburg. *A Syllabus Work-book in Elementary Psychology.* The Macmillan Company, 1930, pp. vii + 110.

Agreeing with Robinson that a college is "a place where there is much teaching, but no learning," the authors have attempted to correct this deficiency by inaugurating a system whereby the student will become the active agent and the professor will become merely the grader of quiz papers. The book consists of very brief introductory statements under each topic, followed by an extensive list of questions bearing upon the subject matter in several standard textbooks. Following these questions is a list of test questions, which the student is to answer. He should then look up the correct answers, given in the appendix, and rate his accomplishment by comparing his score with the appended norms.

The authors confidently expect that the student will attack the *Lessons* with an attitude of mature deliberation; but pessimistic reviewers are likely to believe that students will merely memorize formal answers to a list of questions. The instructions to the student concerning general methods of study, as well as the detailed outline of the course would be invaluable in those colleges where an inadequately prepared instructor finds himself expected to teach beginning psychology. The book would also be useful for any instructor in dealing with those students who have poor study habits. It is doubtful, however, whether these *Developed Lessons* will accomplish with the general student the results which the authors hope to attain.

In contrast to the above, the new book of Remmers and Brandenburg is more of a student's work-book. The pages are all perforated so that the study material may be handed to the instructor upon the completion of the work in each lesson. The first three lessons are given to the discovery of any serious disability on the part of the student so far as reading is concerned and to guide him in the proper study habits for the subject-matter to follow.

A commendable feature, particularly where students are sectioned on the basis of supposed ability, is that the lessons are organized on the two-level plan; that is, the assignments are divided into two parts,

required and optional. Those required are readings from Gates and Dashiell, while the optional assignments cover a wide variety of standard texts and treatises. More valuable, in the reviewers' opinion, would have been the reference to experimental work as it appears in the various journals. On the whole, the work-book exhibits a dearth of experimental citation.

The syllabus work-book is, undoubtedly, superior to many of the mimeographed syllabi of the elementary course which are to be found on practically every campus in the country. Whether or not it can be substituted for these will depend upon agreement among instructors as to the subject-matter and the relative importance of the various topics. For example, in the work-book, eight lessons are devoted to the receiving, connecting, and reacting mechanism; learning, six; motivation, four; native equipment and emotions, three each; and the remaining topics of sensation and feeling, perception and attention, thinking, general intelligence, personality, individuality and efficiency, two each. Few instructors will agree that this distribution of time represents the relative stress which they lay on these topics.

For those who can agree upon the points just mentioned, the *work-book* may be either a material assistance or a serious handicap. If the work is done outside the class-room, then valuable time is saved for a very different presentation during the regularly scheduled hour; but if the syllabus is used as a basis for class discussion, the level of instruction is likely to degenerate.

F. C. Dockeray and Willard L. Valentine,
Ohio State University.

William S. Foster, and Miles A. Tinker. *Experiments in Psychology.* Revised, Henry Holt & Company, 1929, pp. xv + 392.

Norma V. Scheidemann. *Experiments in General Psychology.* University of Chicago Press, 1929, pp. vii + 108.

In the recent revision by Tinker, the long introduction in the earlier Foster has been greatly reduced—as a separate section—and scattered throughout later chapters. Although four experiments of the old edition have been dropped and twelve new ones added so that there are now thirty-two experiments, the actual number is greater than this since several distinct experiments have been included under one heading.

In keeping with the modern trend in psychology, the older experiments on sensation are either omitted or relegated to the precarious status of class demonstrations, while the more social aspects of the science are given more attention. Among these new experiments are "measurement of musical talent," "measurement of art talent," and "measurement of reading ability." Several new experiments on learning are added. New pictures have been added for the experiment in

judging intelligence in photographs. Such factors as sex, background and clothing are kept constant. The appendix also includes comprehensive additions in the way of statistical tables.

On the whole, the book is to be commended. The only serious criticism we have is that the student is given slight opportunity to familiarize himself with psychological apparatus. Aside from an apparatus for recording the extent of the knee-jerk, a color mixer, and a home-made vernier chronoscope, apparatus is conspicuously absent. While laboratories that do not possess equipment will find this a desirable feature, we believe that students gain more respect for the science if adequate apparatus is available for control. The proponents of this latter position will undoubtedly have to face the charge of trying to impress the student; nevertheless, it is our opinion that the manipulation of simple apparatus is a desirable accomplishment for the beginning student.

Scheidemann points out in her preface that the experiments she presents have been "simplified to the point of absurdity," but that the experiment itself is not the goal. She refers to the "experiments" of the "falling apple," the "swinging chandelier," and the "falling weights at Pisa" to support her contention, forgetting that two of these, if not myths are casual observations, and the third was a critical experiment—but three hundred years ago. The use of milk bottle caps for color wheels is an instance of the application of sperm-oil technique in an incandescent age.

If this book were designed for a junior high school, it might have some excuse for existence; but as a serious collection of experiments for the college level, it is surprising that a University Press would sponsor its publication. As for form, the latter book is superior. Foster and Tinker would have been tremendously improved had it been published with large pages and space provided for writing down results.

F. C. Dockeray and Willard L. Valentine,
Ohio State University.

NEW BOOKS AND PAMPHLETS RECEIVED*

Books and pamphlets for review should be sent to James P. Porter, Ohio University, Athens, Ohio.

American Charities and Social Work. Amos Griswold Warner, Stuart Alfred Queen and Ernest Bouldin Harper. Thomas Y. Crowell Company, New York. 616 pp.

L'Attention. H. F. Reynax. 17 Rue Soufflot, Paris. 70 pp.

A Comparative Study of Those Who Accept as Against Those Who Reject Religious Authority. Thomas H. Howells. University of Iowa Studies, Vol. 2, No. 2, University of Iowa, Iowa City, Iowa. 80 pp.

A Comparative Study of Those Who Report the Experience of the Divine Presence and Those Who Do Not. Robert Daniel Sinclair. University of Iowa Studies, Vol. 2, No. 3. University of Iowa, Iowa City, Iowa. 63 pp.

Le Developpement Mental et L'Intelligence. Henri Pieron. Librairie Felix Alcan, Paris, France. 95 pp.

Die Diagnostische Bedeutung der Intelligenzprufungen. Dr. Erich Stern. Verlag Von Theodor Steinkopff, Leipzig, 32 pp.

The Effect of Change of "Drive" on Maze Performance. Merle Hugh Elliott. University of California Press, Berkeley, Calif. 6 pp.

The Effect of the Introduction of Reward Upon the Maze Performance of Rats. Hugh Carlton Blodgett. University of California Press, Berkeley, Calif. 23 pp.

A Further Study in Discrimination of Maze Patterns by the Rat. Joseph G. Yoshioka. University of California Press, Berkeley, Calif. 18 pp.

Kwartalnik Psychologiczny. Stefan Blachowski. Poznan, Dominikakska 4, Poland. 145 pp.

Problems of Neurosis. Alfred Adler. Cosmopolitan Book Corporation, New York City. $3.00. 244 pp.

Social Control of the Mentally Deficient. Stanley P. Davies. Thomas Y. Crowell Company, New York. $3.00. 389 pp.

Survey of College Entrance Credits and College Courses in Music. National Bureau for the Advancement of Music. 45 W. 45th Street, New York. 209 pp.

---

* Mention here does not preclude further comment.

University of Iowa Studies in Psychology. No. XIII. CHRISTIAN A. RUCKMICK, Editor. Psychological Review Company, Princeton, N. J. 214 pp.

Weber's Law in the Discrimination of Maze Distance by the White Rat. JOSEPH G. YOSHIOKA. University of California Press, Berkeley, Calif. 29 pp.

Directory of American Psychological Periodicals (*Continued*)

Journal of Personnel Research—Baltimore: Williams & Wilkins Co. Subscription $5.00. 500 pages approximately. Edited by Walter V. Bingham, and Clarence S. Yoakem. Bi-Monthly.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison and an international co-operating board. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

Psychological Abstracts—Princeton, N. J. Subscription $6.00. 600 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Founded 1925. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Founded 1926. Edited by Edward A. Pace. Published without fixed dates, each number a single research.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $7. 600-700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500-600 pages annually. Edited by John Dewey, Carl Murchison and an international cooperating board. Founded 1929. Quarterly. Political, racial, and differential psychology.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Morrill Hall. Subscription $6.50. 600 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $7. 700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 480 pages annually. Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Founded 1895. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300–400 pp. Founded 1895. Edited by W. S. Hunter. An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116th St. Subscription $4. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge and Wendell T. Bush.

Archives of Psychology—Sub-Station 84, N. Y.: Archives of Psychology. Subscription $6. 500 pp. ann. Founded 1906. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. Subscription $5. Foreign, $5.25. Acting Editor, Henry T. Moore. Quarterly. 432 pages annually. Founded 1906. Abnormal and Social.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $2.50. 288 pages. Ed. by Lightner Witmer. Founded 1907. Without fixed dates (9 Numbers). Orthogenics, psychology, hygiene.

Training School Bulletin—Vineland, N. J.: The Training School. Subscription $1. 160 pp. ann. Ed. by E. R. Johnstone. Founded 1904. Monthly (10 numbers). Psychology and training of defectives.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 500 pages per volume. Edited by W. S. Hunter. Published without fixed dates, each number a single research.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Founded 1916. Bi-monthly. Ed. by Madison Bentley.

Journal of Applied Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 400 pages annually. Founded 1917. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio.

Journal of Comparative Psychology—Baltimore: Williams & Wilkins Co. Subscription $5. 500 pages annually. Founded 1921. Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

(Continued on Third Cover)